Cole - H02573 - 33 - P028980

THE ACCOUNTING REVIEW

VOLUME 80, No. 2 APRIL 2005

P - 229

01.07.05

JENNIFER ALTAMURO, ANNE L. BEATTY, and JOSEPH WEBER
The Effects of Accelerated Revenue Recognition on Earnings Management and Earnings Informativeness: Evidence from SEC Staff Accounting Bulletin No. 101

ORIE E. BARRON, DAVID G. HARRIS, and MARY STANFORD
Evidence That Investors Trade on Private Event-Period Information around Earnings Announcements

LAWRENCE D. BROWN and MARCUS L. CAYLOR
A Temporal Analysis of Quarterly Earnings Thresholds: Propensities and Valuation Consequences

QIANG CHENG and TERRY D. WARFIELD
Equity Incentives and Earnings Management

ANGELA L. COLETTI, KAREN L. SEDATOLE, and KRISTY L. TOWRY
The Effect of Control Systems on Trust and Cooperation in Collaborative Environments

PETER D. EASTON and STEVEN J. MONAHAN
An Evaluation of Accounting-Based Measures of Expected Returns

DAVID B. FARBER
Restoring Trust after Fraud: Does Corporate Governance Matter?

JOSEPH G. FISHER, LAUREEN A. MAINES, SEAN A. PEFFER, and GEOFFREY B. SPRINKLE
An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation

ALOKE GHOSH and DOOCHEOL MOON
Audit Tenure and Perceptions of Audit Quality

JOHN R. M. HAND
The Value Relevance of Financial Statements in the Venture Capital Market

JAYANTHI KRISHNAN
Audit Committee Quality and Internal Control: An Empirical Analysis

XIAOHONG LIU and DAN A. SIMUNIC
Profit Sharing in an Auditing Oligopoly

ANNE M. MAGRO
Knowledge, Adaptivity, and Performance in Tax Research

MOLLY MERCER
The Fleeting Effects of Disclosure Forthcomingness on Management's Reporting Credibility

---

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2004–2005 Executive Committee of the Association are as follows:

President	Jane F. Mutchler, Georgia State University
President-Elect	Judy D. Rayburn, University of Minnesota
Vice President–Education	O. Finley Graves, University of North Texas
Vice President–Finance	Susan F. Haka, Michigan State University
Vice President–International	Lee David Parker, University of Adelaide
Vice President–Professional Relations	Mark M. Chain, Deloitte
Vice President–Publications	Robert Libby, Cornell University
Vice President–Research	John C. Fellingham, The Ohio State University
Vice President–Sections and Regions	Susan V. Crosson, Santa Fe Community College
Past President	William L. Felix, Jr., The University of Arizona

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $135.00	with Electronic Option $155.00
Select two Association-wide journals	U.S. $145.00	with Electronic Option $165.00
All three Association-wide journals	U.S. $155.00	with Electronic Option $175.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $20.00	with Electronic Option $40.00
Select two Association-wide journals	U.S. $30.00	with Electronic Option $50.00
All three Association-wide journals	U.S. $40.00	with Electronic Option $60.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to taruw@u.washington.edu or sent to Professor Terry Shevlin, Senior Editor, *The Accounting Review*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200.

The submission fee of $125.00 for members of the AAA or $200.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826) is published four times a year in January, April, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $170; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to ***The Accounting Review,*** 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2005 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Terry Shevlin
Editors
Lawrence Brown
Patricia Dechow
S. Jane Kennedy
Marlys Gascho Lipe
Madhav Rajan

P28986

33

VOL. 80 APRIL 2005 No. 2

JENNIFER ALTAMURO, ANNE L. BEATTY, and JOSEPH WEBER
The Effects of Accelerated Revenue Recognition on Earnings Management and Earnings Informativeness: Evidence from SEC Staff Accounting Bulletin No. 101 ... 373
ORIE E. BARRON, DAVID G. HARRIS, and MARY STANFORD
Evidence That Investors Trade on Private Event-Period Information around Earnings Announcements 403
LAWRENCE D. BROWN and MARCUS L. CAYLOR
A Temporal Analysis of Quarterly Earnings Thresholds: Propensities and Valuation Consequences 423
QIANG CHENG and TERRY D. WARFIELD
Equity Incentives and Earnings Management 441
ANGELA L. COLETTI, KAREN L. SEDATOLE, and KRISTY L. TOWRY
The Effect of Control Systems on Trust and Cooperation in Collaborative Environments 477
PETER D. EASTON and STEVEN J. MONAHAN
An Evaluation of Accounting-Based Measures of Expected Returns 501
DAVID B. FARBER
Restoring Trust after Fraud: Does Corporate Governance Matter? 539
JOSEPH G. FISHER, LAUREEN A. MAINES, SEAN A. PEFFER, and GEOFFREY B. SPRINKLE
An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation 563
ALOKE GHOSH and DOOCHEOL MOON
Audit Tenure and Perceptions of Audit Quality 585
JOHN R. M. HAND
The Value Relevance of Financial Statements in the Venture Capital Market 613
JAYANTHI KRISHNAN
Audit Committee Quality and Internal Control: An Empirical Analysis 649
XIAOHONG LIU and DAN A. SIMUNIC
Profit Sharing in an Auditing Oligopoly 677
ANNE M. MAGRO
Knowledge, Adaptivity, and Performance in Tax Research 703
MOLLY MERCER
The Fleeting Effects of Disclosure Forthcomingness on Management's Reporting Credibility 723
Editorial Data 745
Editorial Policy and Style Information 746

Editorial Staff of *The Accounting Review* 2004–2005

SENIOR EDITOR
Terry Shevlin, University of Washington
EDITORS
Lawrence Brown, Georgia State University
Patricia Dechow, University of Michigan
S. Jane Kennedy, University of Washington
Marlys Gascho Lipe, University of Oklahoma
Madhav Rajan, Stanford University
EDITORIAL ASSISTANT
Kate Flynn Connolly, University of Washington

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Kathy Casper
Helen Griffith
Judy Cothern
Christine Govaars
Jan Kovarik (subcontracted)
OTHER AAA STAFF
Debora Gardner
Roland LaTulip
Linda Watterworth
Mary Cole
James Szelminski
Barbara Brady
Peggy Turczyn
Myriam Chacon
Deirdre Harris (subcontracted)

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Benjamin C. Ayers, University of Georgia
Frances L. Ayres, University of Oklahoma
William R. Baber, The George Washington University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, Pennsylvania State University
Eli Bartov, New York University
Anne L. Beatty, The Ohio State University
Robert J. Bloomfield, Cornell University
Christine A. Botosan, University of Utah
Robert M. Bowen, University of Washington
Brian J. Bushee, University of Pennsylvania
Robert M. Bushman, The University of North Carolina at Chapel Hill
Paul K. Chaney, Vanderbilt University
C. Bryan Cloyd, University of Illinois at Urbana–Champaign
William M. Cready, The University of Texas at Dallas
Mark L. DeFond, University of Southern California
Michael J. Eames, Santa Clara University
Merle M. Erickson, University of Chicago
Paul Fischer, Pennsylvania State University
Jere R. Francis, University of Missouri–Columbia
Jennifer Francis, Duke University
Jonathan C. Glover, Carnegie Mellon University
Wayne R. Guay, University of Pennsylvania
David A. Guenther, University of Colorado at Boulder
Mahendra Gupta, Washington University
D. Eric Hirst, The University of Texas at Austin
Christopher D. Ittner, University of Pennsylvania
James J. Jiambalvo, University of Washington
Sanjay Kallapur, Purdue University
Ron Kasznik, Stanford University
Bin Ke, Pennsylvania State University
Deen Kemsley, Tulane University
William R. Kinney, Jr., The University of Texas at Austin
Lisa L. Koonce, The University of Texas at Austin
William N. Lanen, University of Michigan
Robert C. Lipe, University of Oklahoma
Joan L. Luft, Michigan State University
Russell J. Lundholm, University of Michigan
Joseph Magliolo III, Southern Methodist University
Laureen A. Maines, Indiana University
Steven R. Matsunaga, University of Oregon
Maureen F. McNichols, Stanford University
Kenneth A. Merchant, University of Southern California
Lillian F. Mills, The University of Arizona
Donald V. Moser, University of Pittsburg
James N. Myers, University of Illinois at Urbana–Champaign
Linda Myers, University of Illinois at Urbana–Champaign
Karen K. Nelson, Rice University
Mark W. Nelson, Cornell University
Patricia C. O'Brien, University of Waterloo
Kathy R. Petroni, Michigan State University
Morton Pincus, The University of Iowa
Jamie H. Pratt, Indiana University
Shivaram Rajgopal, University of Washington
D. Shores, University of Washington
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Theodore Sougiannis, University of Illinois at Urbana–Champaign
Michael T. Stein, University of Oregon
Hun-Tong Tan, Nanyang Technology University/Singapore
Richard M. Tubbs, The University of Iowa
Sandra Vera-Muñoz, University of Notre Dame
Robert E. Verrecchia, University of Pennsylvania
Beverly R. Walther, Northwestern University
Michael Welker, Queen's University
Michael Willenborg, University of Connecticut
Hong Xie, University of Illinois at Urbana–Champaign

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1960–1962	Robert K. Mautz
1963–1964	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William R. Kinney, Jr.
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 373–401

The Effects of Accelerated Revenue Recognition on Earnings Management and Earnings Informativeness: Evidence from SEC Staff Accounting Bulletin No. 101

Jennifer Altamuro
Massachusetts Institute of Technology

Anne L. Beatty
The Ohio State University

Joseph Weber
Massachusetts Institute of Technology

ABSTRACT: The SEC issued Staff Accounting Bulletin (SAB) No. 101 to address its concern that firms were masking true performance by managing earnings using accelerated revenue recognition. Critics of this Accounting Bulletin stated that it would eliminate industry-accepted revenue recognition practices and reduce the quality of reported earnings. The FASB's revenue recognition discussions echo these concerns stating that revenues recorded prior to the completion of the earnings process contain value-relevant information about future performance. This paper investigates these two hypotheses using a sample of firms that accelerated revenue recognition prior to SAB No. 101 adoption (SAB 101 firms) and a matched set of firms that were unaffected by this regulation (unaffected firms). Our earnings distribution tests indicate that SAB 101 firms are more likely to meet earnings benchmarks. Specifically, we find that, in the pre-adoption period, SAB 101 firms report fewer small negative and more small positive earnings than they do in the post-adoption period and than do unaffected firms in the pre-adoption period; SAB 101 firms report fewer small negative and more small positive earnings changes in the pre-adoption period compared to the post-adoption period.

The authors thank Bob Bowen, Mary Barth, Dick Dietrich, S. P. Kothari, Peter Joos, D. Shores, Ram Venkataraman, Charles Wasley, Hal White, Joanna Wu, Jerry Zimmerman, two anonymous reviewers, and seminar participants at Columbia University, University of Florida, The Ohio State University, University of Minnesota, Massachusetts Institute of Technology, Northwestern University, New York University, Pennsylvania State University, Rice University, University of Rochester, Stanford University, and the University of Washington for helpful comments. A previous version of this paper was titled "Motives for Early Revenue Recognition: Evidence from SEC Staff Accounting Bulletin (SAB) 101." Professor Beatty thanks Deloitte & Touche, PricewaterhouseCoopers, and the Sloan Foundation for financial support.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted July 2003
Accepted September 2004

We also document that SAB 101 firms are more likely to have weaker corporate governance and more likely to have financial covenants, providing them with greater incentives to manage earnings. However, we find that the association between earnings and future cash flows and between unexpected earnings and earnings announcement period returns were higher for SAB 101 firms than for unaffected firms in the pre-adoption period, indicating higher earnings informativeness for SAB 101 firms. These associations declined for SAB 101 firms in the post-adoption period, suggesting that SAB No. 101 caused a decline in earnings informativeness. Overall, our results suggest that, although the revenue recognition practices targeted by SAB No. 101 have been used by some firms to manage earnings, the regulation's prohibition of revenue recognition prior to completion of the earnings process, on average, results in less informative earnings since these unearned revenues provide value-relevant information.

Keywords: *revenue recognition; earnings management; benchmark-beating; signaling; cash flows.*

Data Availability: *Data are available from public sources identified in the paper.*

I. INTRODUCTION

The Securities and Exchange Commission (SEC) has argued that improper revenue recognition is the single largest issue contributing to financial restatements.[1] In his famous 1998 speech, "The Numbers Game," Arthur Levitt, then chairman of the SEC, indicated that the use of premature revenue recognition practices to meet Wall Street expectations is one of the five fundamental problems with current financial reporting.[2] Ultimately, the SEC's belief that premature revenue recognition reduces the quality of reported earnings and the integrity of financial reporting caused the SEC to issue Staff Accounting Bulletin (SAB) No. 101, *Revenue Recognition* (SEC 1999).[3]

The SEC states that SAB No. 101 provides guidance on how to interpret and apply existing revenue recognition rules. Critics of SAB No. 101 argue that broad application of those criteria is not always appropriate. For example, Philip Ameen, chair of the Financial Executives Institute Committee on Corporate Reporting, argues that SAB No. 101 changes Generally Accepted Accounting Principles (GAAP), overturning standard industry practices.[4] PricewaterhouseCoopers (2001) further argues that the change in accounting required by SAB No. 101 will result in GAAP earnings that have less predictive ability for future cash flows and thus less relevance in valuing companies.

The Financial Accounting Standards Board (FASB) staff appears sympathetic to the criticisms of SAB No. 101, stating that the regulation goes beyond existing GAAP by making the revenue recognition criteria developed for the software industry the foundation for all basic revenue recognition principles.[5] The FASB also argues that requiring completion of the earnings process for revenue recognition is inconsistent with the recognition criteria for other financial statement components, even though revenues arise from changes in these other components. Based on these concerns and inconsistencies in the Conceptual

---

[1] Turner et al. (2001) identify 381 restatements that occurred during 1997–1999. They find that revenue recognition restatements are the most frequent and result in the largest market reaction.

[2] Levitt's speech, made at the NYU Center for Law and Business on September 28, 1998, can be found at http://www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

[3] See Osterland (2000).

[4] See Ameen (2000) and Osterland (2000) for examples of industry concerns regarding SAB No. 101 adoption.

[5] These arguments can be found in "Project Updates—Revenue Recognition" dated July 14, 2004 at the FASB's website: http://www.fasb.org/project/revenue_recognition.shtml#history.

Framework's treatment of revenue recognition, the FASB has undertaken a project to develop a comprehensive Statement of Financial Accounting Standards on revenue recognition. This standard would go beyond reinstating the revenue recognition practices eliminated by SAB No. 101 by eliminating completeness of the earnings process as a condition for revenue recognition.

In this paper, we use the reporting requirements imposed by SAB No. 101 to examine how accounting methods that accelerate revenue recognition affect financial reporting. First, we examine the SEC's concern that allowing revenue recognition prior to completion of the earnings process results in increased earnings management. Second, we examine the possibility, currently being considered by the FASB, that including unearned revenue in earnings provides value-relevant information about future performance.

To investigate these two hypotheses, we identify a sample of firms that recorded cumulative effects adjustments required by SAB No. 101 adoption (SAB 101 firms) and, where available, we identify the amount of accelerated revenues using the pro forma disclosures. We also identify two sets of control firms whose earnings are unlikely to include accelerated revenue. First, we use the firm as its own control by focusing on the earnings of the SAB 101 firms in the post-adoption period. Second, we use firms that did not report an SAB No. 101 cumulative effects adjustment (unaffected firms) but are in the same industry, have similar asset sizes, and engage in the same types of revenue-generating activities as the SAB 101 firms.

We test our earnings management hypothesis by comparing earnings distributions of SAB 101 firms to earnings distributions of the two sets of control firms to determine whether SAB 101 firms are more likely to beat benchmarks. We also compare characteristics of SAB No. 101 and unaffected firms in the pre-adoption period to determine whether SAB No. 101 firms have stronger incentives to manage earnings.

The results of these tests are consistent with the earnings management hypothesis. Specifically, we find that, in the pre-adoption period, SAB 101 firms report fewer small negative and more small positive earnings than they did in the post-adoption period and than did unaffected firms in the pre-adoption period. These findings suggest that some SAB 101 firms have been accelerating revenue recognition to avoid reporting losses. We also find limited evidence that firms recognize revenue early to avoid reductions in income. In particular, we find that SAB 101 firms report fewer small negative and more small positive earnings changes in the pre-adoption period compared to the post-adoption period, but these differences are not significantly different from those for unaffected firms. Our logistic analysis examining the incentives to manage earnings indicates that SAB 101 firms are more likely to have weak corporate governance and debt with financial covenants. However, we find no evidence that SAB 101 firms are more likely to have recently issued new financing or to have earnings-based bonus plans.

We test the hypothesis that revenues recorded in income prior to the completion of the earnings process are more informative by examining differences in the association between earnings and future cash flows and between unexpected earnings and earnings announcement period returns for our test and control firms. We also use the pro forma data reported by SAB 101 firms to conduct similar tests on the informativeness of accelerated earnings.

The results of these tests are consistent with the earnings informativeness hypothesis. For the SAB 101 firms, the associations between future cash flows, returns, and earnings are higher in the pre-adoption period than in the post-adoption period. When compared to the associations for unaffected firms, the associations for SAB 101 firms are higher in the pre-period and lower in the post-period. We also find that when we use the pro forma information to decompose reported earnings into accelerated and restated earnings, both components predict future cash flows and are associated with contemporaneous annual returns. These findings suggest that, despite our evidence that the revenue recognition practices eliminated by SAB No. 101 were used to manage earnings, on average, these practices resulted in earnings that were more informative.

We link our two sets of results in a supplemental analysis. We partition SAB 101 firms into groups based on their earnings management incentives and examine whether the results of the earnings informativeness tests differs across these groups. We find that, on average, earnings are more highly associated with future cash flows when they contain accelerated revenue recognition regardless of the earnings management incentives. We find mixed results for a similar returns test. Overall, these tests suggest that, on average, earnings that include accelerated revenue are more informative even for groups of firms with the highest incentives to manage earnings.

In summary, the SEC's adoption of SAB No. 101 provides an opportunity to examine the trade-off between discretion in reporting revenues prior to the completion of the earnings process and the relevance of unearned revenues. Although we find some evidence of increased earnings management by firms that accelerate revenue recognition, we find that on average, the relevance of reported earnings for these firms is greater than when the revenue recognition process is delayed. These results are potentially interesting to the regulators and standard setters engaged in developing and implementing revenue recognition standards who must consider the trade-offs between discretion in reporting revenues prior to the completion of the earnings process and the informativeness of unearned revenues.

Section II provides background information for our study. We develop our hypotheses in Section III and discuss our sample selection in Section IV. We provide a description of our research design in Section V. We discuss our empirical results in Section VI and conclude in Section VII.

II. BACKGROUND

In September 1998, Arthur Levitt, then chairman of the SEC, delivered a speech that accused corporate executives of using "accounting hocus-pocus" to "meet or beat Wall Street earnings projections." He argued that manipulation of revenue recognition is one of five popular earnings management tricks employed to "mask the true consequences of management's decisions" (Levitt 1998). In his speech, he challenged "corporate management and Wall Street to re-examine our current environment ... and to embrace nothing less than a cultural change" (Levitt 1998).

In December 1999, the SEC responded to Levitt's challenge by issuing SAB No. 101. The accounting bulletin outlines the SEC's four criteria for revenue recognition: (1) persuasive evidence of an arrangement exists, (2) delivery has occurred or services have been rendered, (3) seller's price to the buyer is fixed or determinable, and (4) collectibility is reasonably assured.

Due to industry concerns about the accounting bulletin, the SEC deferred the effective date of SAB No. 101 until the fourth quarter of 2000. In October 2000, the SEC provided a list of frequently asked questions that stipulated that, consistent with APB No. 20, the transition to SAB No. 101 should be accounted for using a cumulative effects adjustment as long as the prior accounting did not reflect an error.[6] In January 2001, the SEC denied all final requests to change or further delay SAB No. 101 implementation.

Despite the SEC's claims that any group of moderately competent accountants will quickly arrive at the same revenue recognition criteria, in January 2002 the FASB decided to issue, for public comment, a revenue recognition project proposal. The FASB's proposal moves away from the SAB No. 101 earnings process completion requirement for revenue recognition and toward allowing revenue recognition when there is a verifiable measure of a change in the value of a firm's assets or liabilities. The FASB's stated reason for undertaking this project is the lack of a comprehensive standard on revenue recognition and the *ad hoc* nature of implementation guidance. In May 2002, the FASB added this project to its technical agenda and plans to issue a revenue recognition initial due process document in the second quarter of 2005.

III. HYPOTHESIS DEVELOPMENT

The differences between SAB No. 101 and the FASB's revenue recognition proposal highlight the trade-offs in the timing of revenue recognition. The SEC argues that allowing firms to accelerate revenue recognition will allow for misleading earnings management. The FASB's position is that accelerating the recognition of unearned revenues will provide investors with value-relevant information about future performance. This paper provides evidence on these two hypotheses, recognizing that they are not mutually exclusive.

Accelerating revenue recognition could increase the informativeness of earnings in two ways. First, consistent with Subramanyan (1996), some firms choose to manage earnings to signal value-relevant information about future performance. This is contrary to the SEC's views that earnings management is used to mask rather than signal true performance. Second, consistent with the FASB's argument, accelerating revenue recognition increases the value relevance of firms' earnings, on average, even if some firms use accelerated revenue recognition to engage in misleading earnings management.

Earnings Management Hypothesis

Given the SEC's concern, we posit that SAB 101 firms exhibit more evidence of earnings management in the pre-adoption than in the post-adoption period. Burgstahler and Dichev (1997) argue that firms have incentives to just meet or beat earnings benchmarks. Graham et al. (2004) find in a survey of 401 financial executives that over 65 percent of those surveyed thought that having positive earnings is important, 90 percent of those from profitable firms thought meeting earnings change benchmarks is important, and over 40 percent would manipulate revenue recognition polices to meet these benchmarks. If SAB 101 firms were to use the revenue recognition practices targeted by SAB No. 101 to manage earnings, then we would expect these firms to report fewer small negative and more small positive earnings or earnings changes than our control groups.

We also expect that SAB 101 firms have greater incentives to manage earnings than unaffected firms. We consider four earnings management incentives identified by prior research. First, we expect that firms with weak corporate governance are more likely to engage in earnings management. Second, based on Watts and Zimmerman (1986, 1990) we examine the incentive to avoid violating financial covenants and thus reduce the costs of existing debt financing. Third, consistent with Teoh et al. (1998a, 1998b), we examine the incentive to report higher earnings prior to a new capital offering to obtain a better price and increased proceeds from the offering. However, increased SEC scrutiny of public filings for new issuances makes aggressive revenue recognition choices less appealing. The final incentive we consider is CEO bonus plans.

---

[6] APB No. 20 states that the amount of the cumulative effect adjustment shown in the income statement "is the difference between (a) the amount of retained earnings at the beginning of the period of a change and (b) the amount of retained earnings that would have been reported at that date if the new accounting principle had been applied retroactively for all prior periods." APB No. 20 also stipulates that pro forma effects of retroactive application should be shown on the face of the income statement or disclosed prominently in the notes to the financial statements.

Earnings Informativeness Hypothesis

In its current revenue recognition project deliberations, the FASB argues that revenue recognition based on completion of the earnings process is inconsistent with the recognition criteria for other financial statement components, even though revenues arise from changes in these other components. Specifically, it states that the unearned revenues that result from this requirement do not meet the definition of a liability, as defined in Concepts Statement No. 6 (FASB 1985). The FASB has argued that fair values are the most relevant measure of the value of assets and liabilities and changes in these fair values should be the basis for revenue recognition.

This argument is consistent with the views of SAB No. 101 opponents, who state that the bulletin eliminates accepted industry revenue recognition practices that do not reflect an attempt to artificially boost earnings. They contend that the accounting information generated prior to the adoption of SAB No. 101 was more value-relevant because earnings that include accelerated revenue recognition are more useful in predicting future cash flows. For example, if shipments requiring customers' acceptances are not artificially accelerated and instead reflect customer demand, then the amounts associated with these shipments at period end provide an indication of future sales growth.[7]

Both of these arguments suggest that adoption of this Staff Accounting Bulletin affects the association between earnings and future cash flows and between unexpected earnings and earnings announcement period returns for SAB 101 firms. If the SEC is correct, then these associations should increase for SAB 101 firms in the post-adoption period, when firms will no longer be able to accelerate revenue recognition to "mask" declines in performance. If the SAB No. 101 critics are correct, then these associations should decline for SAB 101 firms in the post-adoption period, when firms will no longer be able to use the revenue recognition principles that best reflect the nature of their business.

IV. SAMPLE SELECTION

We use the reporting requirements associated with the enactment of SAB No. 101 to identify a set of firms that elected to accelerate revenue recognition. More specifically, we conduct a keyword search of the Lexis/Nexis SEC filing database, identifying all firms that report a cumulative effect adjustment to earnings resulting from SAB No. 101 adoption. According to the revenue recognition policies proscribed by SAB No. 101, each of these firms accelerated revenue recognition prior to the adoption of this standard. This search yields a sample of 257 firms that adjust earnings. Compustat data was unavailable for 28 of these firms, resulting in a final sample of 229 SAB 101 firms. The two most common types of accelerated revenue recognition polices in our sample relate to upfront fees under license agreements and shipment of goods requiring customer acceptance. The Appendix provides examples of the financial statement disclosures associated with these types of revenue polices before and after SAB No. 101 adoption.

---

[7] See Rajgopal et al. (2003) for a similar argument regarding the value relevance of order backlog. The authors find that backlog amounts reported in the financial statements do provide information about future earnings. However, their results indicate that the stock market behaves as if backlog contributes more to future earnings than that implied by the cross-sectional association between backlog and future earnings.

Only 110 of our SAB 101 firms (48 percent) reported pro forma information related to the adjustment, even though APB No. 20 requires the disclosure of this information. If available, the pro forma data provides a measure of as-if earnings for our SAB 101 firms. Given the reduction in our sample that results from using only this data and the potential for selection bias, we identify two sets of control firms whose earnings are unlikely to be affected by accelerated revenue recognition as an alternative method of measuring as-if earnings.

Our first set of control firms is our SAB 101 firms in the post-adoption period. We argue that in the pre-period the firms affected by SAB No. 101 were accelerating revenue recognition, but that in the post-period they were not. The advantage of using a firm as its own control is that the test sample and control sample are likely to be similar along all dimensions other than the revenue recognition policy followed. The disadvantage is that tests spanning the economic downturn in 2000, the year firms adopted SAB No. 101, could potentially bias our tests toward finding a change in the propensity to meet benchmarks and a change in the value relevance of earnings.

Our second control sample consists of firms unaffected by SAB No. 101, but otherwise similar to the SAB 101 firms. To identify these firms, we first match our SAB 101 firms to other Compustat firms based on their two-digit SIC code and asset size. We then verify that SAB 101 firms and unaffected firms engage in the same types of revenue-generating transactions. Specifically, we examine whether our unaffected firms report engaging in economic activity that generates upfront fees and customer acceptances, which are the two most frequent types of activities affected by this accounting change.[8] We determine this by performing a keyword search on the financial statements of our unaffected firms. Despite substantial differences in how firms discuss their operations, we find a keyword match for 89 percent of the sample and identify 226 unaffected control firms.

By employing these two control samples, we not only compare our SAB 101 firms to unaffected firms in the pre-adoption period, but we also incorporate difference-in-differences tests that compare changes in test firms relative to changes in control firms. This differences-in-differences research design helps to mitigate the potential impact of the economic downturn on our tests.[9]

V. RESEARCH DESIGN

Earnings Management Tests

To provide evidence on our earnings management hypothesis, we examine the earnings distributions of SAB 101 firms and unaffected firms to test for differences in the propensity to meet benchmarks. We use Degeorge et al.'s (1999) hierarchy of benchmark importance. We first examine the likelihood of meeting a positive earnings benchmark. Conditional on the firms having positive earnings, we examine the likelihood of meeting an earnings change benchmark.[10]

---

[8] Of the 229 sample firms, 50 firms adjusted their financial statements for other reasons. For these firms we examined whether the adjusting and unaffected firms were in the same three- or four-digit SIC code. We find that 37 percent are in the same three-digit SIC code and 11 percent are in the same four-digit SIC code.

[9] The effectiveness of the design in reducing this problem depends on the extent to which the economic downturn affected adjusting and unaffected firms similarly. To determine the relative impact, we examine the effect of the downturn on the earnings distribution of these two samples. We find that the economic downturn had a similar impact on the 1st quartile, median, and 3rd quartile firms in the adjusting and unaffected samples.

These benchmark tests use quarterly data. We include all quarters from fiscal 1997–1999 in our pre-adoption period. We exclude the four quarters of fiscal 2000 from our comparisons because it is unclear whether they reflect the implementation of SAB No. 101. The post-adoption period includes the eight fiscal quarters in 2001 and 2002, and any fiscal quarters with data available in fiscal 2003.

Our earnings-level tests use quarterly net income (Compustat data item 69) scaled by end-of-period total assets (Compustat data item 44).[11] Our earnings change analysis uses the seasonally adjusted difference in quarterly net income scaled by end-of-period total assets.[12] In both our earnings levels and earnings changes analyses, we define the distribution bin width to be 0.75 percent. We classify firms in the first bin to the right of zero as just meeting the benchmark, while those in the first bin to the left of zero are classified as just missing the benchmark.[13]

For the SAB 101 and unaffected firm samples, we create pre-adoption period and post-adoption period earnings histograms. Thus, we create four histograms of quarterly earnings and four histograms for changes in quarterly earnings. Our tests require that we develop a model of the number of firms expected to fall in each bin in the absence of earnings management. For each histogram, we calculate this expectation using the Burgstahler and Dichev (1997) method of averaging the number of observations in the bins to the right and to the left of the bin of interest. We then create a variable that measures the difference between the actual number of firms and the expected number of firms for each bin in each earnings and earnings change histogram (*Earn_Diff*, *Earnchng_Diff*, respectively). To measure the extent of earnings management to meet benchmarks we create an indicator variable, *Netbin*, that is 1 for the bin just to the right of a particular benchmark (i.e., zero earnings or same quarter last year's earnings), −1 for the bin just to the left of the benchmark, and 0 otherwise.

We estimate the following regressions for the SAB 101 firms, the unaffected firms, and the pooled sample to test whether firms that accelerated revenue recognition were more likely to meet earnings thresholds:

---

[10] We do not examine the analyst forecast benchmark, the third component of Degeorge et al.'s (1999) hierarchy, because we do not feel that measurement of a firm's ability to meet this benchmark is applicable to our setting. In the pre-period, the ongoing communication between managers and analysts makes it difficult to form inferences about what information analysts had about accelerated revenue and the ability of firms to use revenue recognition policies to meet the analyst forecast benchmarks. Additionally, the adoption period for Reg FD and SAB No. 101 overlap. Evidence suggests Reg FD has affected managers' ability to guide analysts' forecasts, and thus the firms' propensity to meet analyst benchmarks.

[11] We focus on net income rather than other earnings measures such as pre-tax income based on Burgstahler's (2004) discussion of Beaver et al. (2003). Specifically, he comments that bottom line earnings are likely to be the focus of manager's benchmark-beating behavior. The most appropriate scalar was not obvious to us, so we rerun all of our tests scaling by the end-of-period shares outstanding (quarterly Compustat data item 61) and by market value of equity (quarterly Compustat data item 61 multiplied by quarterly Compustat data item 14) and find qualitatively similar results.

[12] Graham et al.'s (2004) survey results suggest that CEOs focus more on the seasonally adjusted change in net income than on the change in net income in adjacent quarters. We find similar results using this alternative earnings change metric.

[13] Since bin width is an arbitrary research design choice, we investigate the sensitivity of our results to this choice. Using bin widths of 0.5 percent to 1 percent, we find qualitatively similar results.

$$Earn\_Diff_{(b,t)} = \alpha + \beta_1 Post + \beta_2 Netbin_{(b,t)} + \beta_3 Post{*}NetBin_{(b,t)} + \varepsilon \tag{1}$$

$$Earnchng\_Diff_{(b,t)} = \alpha + \beta_1 Post + \beta_2 Netbin_{(b,t)} + \beta_3 Post{*}NetBin_{(b,t)} + \varepsilon \tag{2}$$

where:

$Earn\_Diff_{(b,t)}$ = difference between the expected number of firms and the actual number of firms for each bin *b* in each period *t* for the histogram of net income (quarterly Compustat data item 69) scaled by end-of-period assets (quarterly Compustat data item 44); the bin width is 0.75 percent; the expected number of firms for each bin *b* in each period *t* is estimated using the average of the number of observations in the adjacent bins to the left and to the right of the bin of interest;

$Earnchng\_Diff_{(b,t)}$ = difference between the expected number of firms and the actual number of firms for each bin *b* in each period *t* for the histogram of the seasonally adjusted change in net income (quarterly Compustat data item 69) scaled by the end-of-period assets (quarterly Compustat data item 44); the bin width is 0.75 percent; the expected number of firms for each bin *b* in each period *t* is estimated using the average of the number of observations in the adjacent bins to the left and to the right of the bin of interest.

$Netbin_{(b,t)}$ = indicator variable equal to 1 if the firm falls in the histogram bin just to the right of the benchmark, −1 for the histogram bin just to the left of the benchmark, and 0 otherwise;

Post = indicator variable equal to 1 if the observation is drawn from the post-adoption period earnings distribution, and 0 otherwise; and

$Post{*}Netbin_{(b,t)}$ = calculated by multiplying *Post* by *Netbin.*

We conduct t-tests on the coefficients on *Post*Netbin* to determine whether the propensity to meet thresholds was affected by SAB No. 101 adoption. Using an interactive dummy variable, we conduct F-tests on the differences between the coefficients for SAB 101 and unaffected firms on *Netbin* to test for significant differences between these two samples in the propensity to meet benchmarks in the post-adoption period. We also conduct F-tests on *Post*Netbin* to test for significant differences in the changes in the propensity to meet benchmarks for the pre- versus post-adoption period across these two samples.

We further test our earnings management hypothesis by examining firms' earnings management incentives. Specifically, we compare the characteristics of the SAB 101 firms to the unaffected firms during the pre-adoption period.

To proxy for the quality of a firm's corporate governance system, we focus on the number of shares held by outside members of the board of directors as reported in the firm's fiscal year 1999 proxy statement. Consistent with Bushman et al. (2002), we classify a member of the board as an outsider if he or she is not a member of the management team, a retired member of the management team, or a family member of the management team of the firm. We then create a variable, *Corp Gov*, which is the sample rank of the percentage of the shares owned by the board of directors that is owned by outside members.

We focus on the firm's use of bank debt financing to proxy for the debt-contracting incentives to manage earnings. Begley and Freedman (2004) document that in the period 1999–2000, public debentures do not include financial covenants. Since bank debt virtually always includes financial covenants, we argue that firms with bank debt are more likely to be affected by financial covenants. Accordingly, we create an indicator variable, ***Bank Debt***, which is 1 if the firm has bank debt, and 0 otherwise. We examine the firms' debt footnotes and the material contracts section of their 1999 10-Ks to determine if they have outstanding bank debt.

We use the Securities Datacorp Database (SDC) of debt and equity issuances to create three variables that proxy for the incentives to manage earnings prior to the issuance of debt or equity. We measure these variables during the three-year period prior to SAB No. 101 adoption. The first, ***IPO***, is an indicator variable equal to 1 if the firm undertook an IPO during the period, and 0 otherwise. The second, ***Equity Issue***, is an indicator variable equal to 1 if the firm issued equity during the period, and 0 otherwise. The third, ***Debt Issue***, is an indicator variable equal to 1 if the firm issued public debt during the period, and 0 otherwise.

To measure the effect of CEO bonuses as an earnings management incentive, we examine fiscal year 1999 proxy statements to determine whether firms' bonus plans are based on financial performance measures. Managers of firms with accounting-based bonus plans are expected to have stronger incentives to manage earnings upwards. Our proxy for this incentive, ***Bonus***, is an indicator variable equal to 1 if the firm's bonus plan is based on financial performance, and 0 otherwise. Finally, we also include *Size* as a control variable, measured as the log of assets (annual Compustat data item 6) at the end of the pre-adoption period. Using these proxies we run the following logistic regression:

$$\begin{aligned} SAB = {} & \alpha + \beta_1 Corp\ Gov + \beta_2 Bank\ Debt + \beta_3 IPO + \beta_4 Equity\ Issue \\ & + \beta_5 Debt\ Issue + \beta_6 Bonus + \beta_7 Size + \varepsilon \end{aligned} \tag{3}$$

where:

SAB = 1 for SAB 101 firms, and 0 for unaffected firms;
Corp Gov = sample rank of percentage of board of director ownership shares that are held by outside directors;
Bank Debt = equals 1 if the firm has outstanding bank debt, and 0 otherwise;
IPO = 1 if the firm undertook an initial public offering during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;
Equity Issue = 1 if the firm issued equity during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;
Debt Issue = 1 if the firm's issued public debt during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;
Bonus = 1 if the firm offered a cash bonus based on financial performance in the fiscal year prior to SAB No. 101 adoption, and 0 otherwise; and
Size = the log of assets (annual Compustat data item 6) at the end of the fiscal year prior to SAB No. 101 adoption.

Earnings Informativeness Tests

We use two measures of performance to test our earnings informativeness hypothesis. Our first construct is future cash flows. In the Statement of Concepts, the FASB argues that "financial reporting should provide information to help investors, creditors, and others assess the amounts, timing, and uncertainty of prospective net cash inflows to the related enterprise."[14] Based on the FASB's argument, we assert that earnings that are more highly associated with future cash flows are more value-relevant.[15] Our second construct is the market's response to a firm's earnings announcement. We argue that if investors perceive earnings to be more informative, then there will be higher earnings response coefficients (ERCs).[16] Consistent with distinctions made in Lev (1989), we choose a short window during which to examine the association between earnings and returns to assure that the price change is mostly due to earnings information, rather than examining a longer window that might include other information and lead us to overstate the incremental information contribution of earnings to returns.

We estimate the following regressions controlling for fixed firm and quarter effects, for the SAB 101 firms, the unaffected firms, and the pooled sample to test whether accelerated revenue recognition results in earnings that are more informative:[17]

$$LeadCFO_{(i,q+t)} = \alpha + \beta_1 Post + \beta_2 Earn_{(i,q)} + \beta_3 Post{*}Earn_{(i,q)} + \varepsilon \tag{4}$$

$$Returns_{(i,t)} = \alpha + \beta_1 Post + \beta_2 \Delta Earn_{(i,q)} + \beta_3 Post{*}\Delta Earn_{(i,q)} + \varepsilon \tag{5}$$

where:

$LeadCFO_{(i,q+t)}$ = accumulated cash flow from operations (quarterly Compustat data item 108) scaled by end-of-period assets (quarterly Compustat data item 44) for firm *i* in one, two, three, and four subsequent quarters;
$Returns_{(i,t)}$ = two-day cumulative market-adjusted returns from CRSP (using the value weighted market index) for firm *i* on the earnings report date *t* and date *t*−1;
Post = indicator variable equal to 1 if the firm-quarter observation occurs during the post-adoption period, and 0 otherwise;
$Earn_{(i,q)}$ = earnings, measured as the net income (quarterly Compustat data item 69) for firm *i* in period *q*, scaled by end-of-period assets (quarterly Compustat data item 44);
$Post{*}Earn_{(i,q)}$ = *Post* multiplied by *Earn*;
$\Delta Earn_{(i,q)}$ = unexpected earnings, measured as the seasonally adjusted change in net income (quarterly Compustat data item 69) for firm *i* in period *q*, scaled by end-of-period shares outstanding (quarterly Compustat data item 61); and
$Post{*}\Delta Earn$ = *Post* multiplied by $\Delta Earn$.

We conduct t-tests on the coefficients on *Post*Earn* and *Post*ΔEarn* to evaluate whether the change in revenue recognition practices resulting from SAB No. 101 adoption significantly affected earnings informativeness. We conduct F-tests on the differences between the coefficients for SAB 101 and unaffected firms on *Earn* and *ΔEarn* to test for a significant difference between the two samples in earnings informativeness in the pre-adoption period. We also conduct an F-test on *Post*Earn* and *Post*ΔEarn* to test for significant differences in the changes in informativeness for the pre-adoption versus post-adoption periods across these two samples of firms. Like our earnings management tests, these tests use fiscal 1997–1999 quarterly data for the pre-adoption period and use the eight fiscal quarters in 2001 and 2002, and any fiscal quarters with data available in fiscal 2003 for our post-adoption period.

---

[14] See SFAC No. 1, FASB (1978, para. 37).

[15] See Dechow et al. (1998) and Barth et al. (2001) for similar arguments on why the association between reported earnings and future cash flows is a good measure of earnings informativeness.

[16] See Lev (1989) and Kothari (2001) for a summary of the extensive research examining the association between a firm's earnings and stock returns.

[17] We choose scalars for each performance metric based on previous research and scale the independent variables with the same scalar as the dependent variable. Consequently, earnings and lead cash flows are both scaled by assets in our cash flow analysis, while unexpected earnings are scaled by shares outstanding. As we discuss below, we examine the sensitivity of our results to alternative scalars.

We perform a second set of informativeness tests using the pro forma data collected from the SAB 101 firm's fiscal year 2000 10-K filing to measure as-if earnings for periods reported prior to the adjustment. More specifically, we use the pro forma reconciliation to separate earnings into two parts: (1) earnings that do not include accelerated revenue recognition (*Pro Forma Earn*), and (2) earnings as a result of accelerated revenue (*Accelerated Earn*). Firms that chose to report pro forma earnings presented this information for either fiscal years 1998 and 1999, or fiscal year 1999 only. Since the pro forma data is only reported for annual earnings, we regress one-year-ahead cash flows and annual returns on *Pro Forma Earn* and *Accelerated Earn* for each year provided in the reconciliation to investigate whether the earnings associated with the acceleration of revenue have a statistically different association with future annual cash flows or annual returns. Our annual return accumulation period begins at the end of the fourth month of fiscal year t and ends at the end of the fourth month of fiscal year $t+1$. We estimate the following regressions:

$$Annual\ LeadCFO_{(i,t+1)} = \alpha + \beta_1 Pro\ Forma\ Earn_{(i,t)} + \beta_2 Accelerated\ Earn_{(i,t)} + \varepsilon \quad (6)$$

$$Annual\ Returns_{(i,t)} = \alpha + \beta_1 Pro\ Forma\ Earn_{(i,t)} + \beta_2 Accelerated\ Earn_{(i,t)} + \varepsilon \quad (7)$$

where:

$Annual\ LeadCFO_{(i,t+1)}$ = annual cash flow from operations (annual Compustat data item 308) scaled by end-of-period assets (annual Compustat data item 6) for firm i in year $t+1$;

$Annual\ Returns_{(i,t)}$ = annual cumulative market-adjusted returns from CRSP (using the value-weighted market return) for firm i from end of month four of fiscal year t until the end of month four of fiscal year $t+1$;

$Pro\ Forma\ Earn_{(i,t)}$ = difference between actual earnings reported for firm i in period t and the amount of the pro forma adjustment associated with a cumulative effect from an accounting change related to SAB No. 101, scaled by end-of-period assets; and

$Accelerated\ Earn_{(i,t)}$ = amount of earnings related to periods after fiscal year 2000 that had been recognized in a prior period earnings (i.e., earnings that had been accelerated in periods prior to fiscal year 2000) scaled by end-of-period assets.

VI. RESULTS

Table 1 provides descriptive data by industry for all industries with five or more SAB 101 firms. Specifically, we report the average adjustment size scaled by market value of equity, the types of adjustments made in each industry and the total number of adjustments in the industry.

TABLE 1
Average Industry Restatement Amount, Scaled by Market Value of Equity and Number of SAB 101 Firms by Type of Adjustment

Industry	SIC Code	Restatement/ MVE	Number of Up-Front Fees	Number of Customer Acceptances	Other Changes in Revenue Recognition	Total Number of Adjustments
Food	20	0.001	0	5	0	5
Pharmaceuticals and Chemicals	28	0.035	41	2	0	43
Machinery	35	0.024	1	15	1	17
Electrical	36	0.008	2	6	0	8
Measuring Instruments	38	0.023	9	18	1	28
Communications	48	0.005	18	2	0	20
Clothing Retail	56	0.001	0	0	5	5
Miscellaneous Retail	59	0.006	1	6	2	9
Holding Companies	67	0.008	2	0	5	7
Computer and Business Services	73	0.029	17	8	7	32
Management Services	87	0.052	4	0	3	7
All Other			5	17	26	48
Total		0.029	100	79	50	229

The average earnings adjustment in our sample is nearly 3 percent of the market value of equity (MVE), while the largest average adjustment, 5.2 percent of MVE, is in the Management Services industry. We also find that the types of adjustments are clustered by industry. Specifically, non-refundable payments were most commonly recognized in the Pharmaceutical and Chemical, Communications, and Computer and Business Services industries. In the Machinery and Measuring Instruments industries, it was more common to recognize revenue prior to final customer acceptance. The largest number of adjustments, 43, is in the Pharmaceuticals and Chemicals industry. We also find that our sample is heavily concentrated in the Computer and Business Services and Measuring Instruments industries.

Earnings Management Results

Earnings Benchmarks

Table 2 presents the results of estimating models (1) and (2), our earnings benchmark regressions. Panel A reports the results for earnings levels. For SAB 101 firms, the statistically significant positive coefficient on *Netbin* indicates that, in the pre-adoption period, SAB 101 firms are more likely to have small positive earnings and less likely to have small negative earnings. The statistically significant negative coefficient on *Post*Netbin* indicates that the discontinuity in the earnings distribution declines from the pre- to the post-adoption period.

The coefficient on *Netbin* for unaffected firms is also positive and statistically significant. This indicates that in the pre-adoption period, unaffected firms also report more small positive earnings and fewer small negative earnings. In contrast to the findings for SAB 101 firms, we find no evidence that this likelihood declined in the post-adoption period for unaffected firms.

In the third column, we report the results for our pooled sample regression testing for a statistical difference in the SAB 101 and unaffected firms' propensity to meet benchmarks. The results indicate that in the pre-adoption period SAB 101 firms were more likely to meet benchmarks than unaffected firms. Furthermore, the reduction in the propensity to meet benchmarks associated with the adoption of SAB No. 101 affects SAB 101 firms more than unaffected firms. Together, these results are consistent with the hypothesis that SAB 101 firms were managing earnings in the pre-adoption period by accelerating revenue recognition to report fewer small losses and more small profits.

Panel B of Table 2 reports the results from a similar analysis using the distribution of earnings changes for firms with positive net income. Similar to earnings levels, in the pre-adoption period, SAB 101 firms report more small positive and less small negative earnings changes than in the post-adoption period. We also find evidence that unaffected firms were more likely to meet the positive earnings change benchmark in the pre-adoption period; however, we find no evidence that this likelihood changes in the post-period. The tests of differences between the SAB 101 firms and the unaffected firms are not statistically significant. Overall, this analysis provides limited evidence that SAB 101 firms accelerated revenue recognition to meet the positive earnings change benchmark.[18]

---

[18] To ensure that our results are not driven by our method of calculating the expected number of observations in a bin, we repeat the analysis using an alternative method. We use nonlinear least squares to construct a curve that fits the histogram of each of the earnings distributions. The estimated function is used to calculate the expected number of firms for each bin. When we use this expectation, the results of our tests (not tabulated here) are qualitatively similar.

Incentives to Manage Earnings

Table 3 presents univariate tests for differences in means between SAB 101 and unaffected firms for the earnings management incentives proxies. Due to data availability, we compare firm characteristics of 187 SAB 101 firms and 158 unaffected firms. We find SAB 101 firms have weaker corporate governance (as indicated by less outsider share ownership) than unaffected firms, and 73 percent of SAB 101 firms have bank debt compared to 61 percent of unaffected firms. We do not note a significant difference in the issuance of debt or equity securities in the three years prior to SAB No. 101 adoption, with the exception

---

TABLE 2

Coefficients and (t-statistics) from a Regression of the Extent of the Discontinuity in the Earnings Distribution for a Sample of SAB 101 Firms and a Sample of Unaffected Firms before and after the Effective Date of SAB No. 101

Panel A: Earnings Levels

$$Earn\_Diff_{(b,t)} = \alpha + \beta_1 Post + \beta_2 Netbin_{(b,t)} + \beta_3 Post{*}NetBin_{(b,t)} + \varepsilon$$

Variable	Predicted Sign	SAB 101 Firms Coefficients (t-statistics)	Unaffected Firms Coefficients (t-statistics)	Difference in Coefficients (F-test)
Intercept	+/−	−0.02	0.02	−0.04
		(−0.02)	(0.02)	(−0.03)
Post	+/−	0.02	0.05	−0.03
		(0.02)	(0.03)	(−0.01)
Netbin	+/−	68.625	40.75	27.88
		(14.99)***	(10.53)***	(4.65)***
*Post*Netbin*	+/−	−40.625	−1.75	−38.89
		(−6.27)***	(−0.32)	(−4.59)***
Number of Obs.[a]		139	139	
Adjusted R^2		0.66	0.60	

Panel B: Earnings Changes

$$Earnchng\_Diff_{(b,t)} = \alpha + \beta_1 Post + \beta_2 Netbin_{(b,t)} + \beta_3 Post{*}NetBin_{(b,t)} + \varepsilon$$

Variable	Predicted Sign	SAB 101 Firms Coefficients (t-statistics)	Unaffected Firms Coefficients (t-statistics)	Difference in Coefficients (F-test)
Intercept	+/−	−0.01	−0.01	−0.01
		(−0.00)	(−0.00)	(−0.00)
Post	+/−	0.01	0.01	0.01
		(0.01)	(0.01)	(0.01)
Netbin	+/−	29.125	19.75	9.375
		(3.15)***	(1.55)*	(0.60)
*Post*Netbin*	+/−	−35.875	−17.50	−18.375
		(−2.74)**	(−0.97)	(−0.86)
Number of Obs.[a]		115	115	
Adjusted R^2		0.05	0.01	0.16

(*continued on next page*)

TABLE 2 (continued)

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

[a] Represents the number of histogram bins measured in the pre-period plus the number of bins measured in the post-period for each distribution type (earnings, earnings changes) and for each sample (SAB 101 and unaffected).

Variable Definitions:

Earn_Diff$_{(b,t)}$ = difference between the expected number of firms and the actual number of firms for each bin *b* in each period *t* for the histogram of net income (quarterly Compustat data item 69) scaled by end-of-period assets (quarterly Compustat data item 44). The bin width is 0.75 percent. The expected number of firms for each bin *b* in each period *t* is estimated using the average of the number of observations in the adjacent bins to the left and to the right of the bin of interest;

Earnchng_DIFF$_{(b,t)}$ = difference between the expected number of firms and the actual number of firms for each bin *b* for each period *t* for the histogram of the seasonally adjusted change in net income (quarterly Compustat data item 69) scaled by the end-of-period assets (quarterly Compustat data item 44). The bin width is 0.75 percent. The expected number of firms for each bin *b* in each period *t* is estimated using the average of the number of observations in the adjacent bins to the left and to the right of the bin of interest;

Netbin$_{(b,t)}$ = indicator variable equal to 1 if the firm falls in the histogram bin just to the right of the benchmark, −1 for the histogram bin just to the left of the benchmark, and 0 otherwise;

Post = indicator variable equal to 1 if the observation is drawn from the post-adoption period earnings distribution, and 0 otherwise; and

*Post*Netbin*$_{(b,t)}$ = calculated by multiplying *Post* by *Netbin*.

of IPO offerings. We find that only 9 percent of the SAB 101 firms issued IPOs, compared to 15 percent of the unaffected firms. We find no difference in the use of financial performance bonuses in the two samples. Also, there is no significant difference in asset size for our two samples, indicating that our size match was effective.

Table 4 presents the results of estimating model (3), our logistic regression comparing the earnings management incentives of SAB 101 and unaffected firms. We find that SAB 101 firms have weaker corporate governance than unaffected firms. This finding suggests that firms with good corporate governance adopt more conservative accounting policies. We also find that SAB 101 firms are more likely to have bank debt. This result is consistent with financial covenants affecting the revenue recognition decision.

We find no differences in the propensity of SAB 101 and unaffected firms to issue debt or equity securities or to go public. We also do not find a difference between the two samples in the use of bonuses based on financial performance. This is consistent with the evidence in Graham et al. (2004) that financial executives consider bonuses to be a relatively less important earnings management incentive.

Earnings Informativeness Results

Cash Flow Analysis

Table 5 presents the results of estimating model (4), a regression of future operating cash flows on current earnings. Panel A reports the association between quarter *q* earnings and one-quarter-ahead cash flows. Panels B, C, and D report the results of similar analyses using future cash flows accumulated for two quarters, three quarters, and four quarters, respectively.

In Panel A of Table 5 we find a statistically significant positive coefficient on *Earn* for SAB 101 firms indicating a positive association between earnings and one-quarter-ahead cash flows in the pre-adoption period. A decline in the association between earnings and

TABLE 3
Univariate Tests of Differences in Mean Values of Firm Characteristics for a Sample of SAB 101 Firms and a Sample of Unaffected Firms Matched on Size and Industry Membership

Variable	Predicted Sign of Difference	Mean for SAB 101 Firms	Mean for Unaffected Firms	Difference in Means	t-statistic of Difference
Corp Gov	−	178.80	201.70	−22.90	−1.99**
Bank Debt	+	0.73	0.61	0.12	2.33**
IPO	+/−	0.09	0.15	−0.06	−1.80*
Debt Issue	+/−	0.14	0.18	−0.04	−1.03
Equity Issue	+/−	0.38	0.31	0.07	1.38
Bonus	+	0.81	0.84	0.03	−0.70
Size	+/−	5.51	5.56	−0.05	−0.21
Num Obs.		187	158		

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

Variable Definitions:

Corp Gov = sample rank of percentage of board of director ownership shares that are held by outside directors;
Bank Debt = 1 if the firm has outstanding bank debt, and 0 otherwise;
IPO = 1 if the firm undertook an initial public offering during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;
Equity Issue = 1 if the firm issued equity during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;
Debt = 1 if the firm's issued public debt during the three-year period prior to SAB No. 101 adoption, and 0 otherwise.
Bonus = 1 if the firm offered a cash bonus based on financial performance in the fiscal year prior to SAB No. 101 adoption, and 0 otherwise; and
Size = the log of assets (annual Compustat data item 6) at the end of the fiscal year prior to SAB No. 101 adoption.

future cash flows in the post-adoption period is indicated by the negative coefficient on *Post*Earn*. This is consistent with the FASB's concern that SAB No. 101 has led to a decline in earnings informativeness.

The second column of Panel A of Table 5 indicates that there is also a positive correlation between earnings and future cash flows for unaffected firms in the pre-adoption period. In contrast to SAB 101 firms, unaffected firms show no change in the association between earnings and future cash flows in the post-adoption period. F-tests of differences between the SAB 101 and unaffected firms indicate that, in the pre-adoption period, SAB 101 firms' earnings were more informative than earnings of the unaffected firms. In the post-adoption period, SAB 101 firms had a statistically significant reduction in the association of earnings and future cash flows when compared to the unaffected firms' association.

We find a similar pattern in Panels B and C for two-quarter-ahead and three-quarter-ahead cumulative cash flows. The only difference in significance between these two panels and Panel A is the lack of significance in Panel C on the reduction in the association for SAB 101 firms in the post-adoption period. However, the results in Panels B and C do indicate that the incremental informativeness of earnings for SAB 101 firms, compared to unaffected firms, deteriorates as the period of accumulation of cash flows is lengthened. In

TABLE 4

Coefficients and (Chi-Squared Statistics) from a Logistic Regression Comparing the Earnings Management Incentives of a Sample of SAB 101 Firms and a Sample of Unaffected Firms Matched on Size and Industry Membership

$$SAB = \alpha + \beta_1 Corp\ Gov + \beta_2 Bank\ Debt + \beta_3 IPO + \beta_4 Equity\ Issue + \beta_5 Debt\ Issue + \beta_6 Bonus + \beta_7 Size + \varepsilon$$

Variable	Predicted Sign	Coefficient (Chi-Squared)
Intercept	+/−	0.53 (1.46)
Corp Gov	−	−0.002 (−4.10)**
Bank Debt	+	0.66 (7.01)***
IPO	+/−	−0.51 (−2.13)
Equity Issue	+/−	0.02 (0.01)
Debt Issue	+/−	−0.27 (−0.58)
Bonus	+	−0.27 (−0.85)
Size	+/−	−0.02 (−0.06)
Number of Obs.		345
Pseudo R^2		3.5%
Percent Correctly Predicted		61.8%

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

Variable Definitions:

SAB = 1 for SAB 101 firms, 0 for unaffected firms;

Corp Gov = sample rank of percentage of board of director ownership shares that are held by outside directors;

Bank Debt = 1 if the firm has outstanding bank debt, and 0 otherwise;

IPO = 1 if the firm undertook an initial public offering during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;

Equity Issue = 1 if the firm issued equity during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;

Debt Issue = 1 if the firm's issued public debt during the three-year period prior to SAB No. 101 adoption, and 0 otherwise;

Bonus = 1 if the firm offered a cash bonus based on financial performance in the fiscal year prior to SAB No. 101 adoption, and 0 otherwise; and

Size = the log of assets (annual Compustat data item 6) at the end of the fiscal year prior to SAB No. 101 adoption.

Panel D, we find that in the pre-adoption period, SAB 101 firms' earnings are more informative than unaffected firms, but SAB No. 101 adoption does not have a marginal impact on SAB 101 firms.

Collectively, these results demonstrate that accelerating revenue recognition results in earnings that are more informative about future cash flows. Restrictions on the acceleration

TABLE 5

Coefficients (t-statistics) from a Firm and Quarterly Fixed-Effects Regression of the Association between Earnings and Cash Flows Accumulated One Quarter Ahead, Two Quarters Ahead, Three Quarters Ahead, and Four Quarters Ahead for a Sample of SAB 101 Firms and a Sample of Unaffected Firms

$$LeadCFO_{(i,q+t)} = \alpha + \beta_1 Post + \beta_2 Earn_{(i,q)} + \beta_3 Post{*}Earn_{(i,q)} + \varepsilon$$

Variable	Predicted Sign	SAB 101 Firms Coefficients (t-statistics)	Unaffected Firms Coefficients (t-statistics)	Difference in Coefficients (F-test)
Panel A: *Cash Flows*$_{q+1}$				
Intercept	+/−	−0.001	0.002	−0.003
		(−0.29)	(1.22)	(−1.06)
Post	+/−	0.001	−0.008	0.009
		(0.15)	(−2.04)**	(1.55)
Earn	+/−	0.45	0.23	0.22
		(17.35)***	(10.87)***	(6.79)***
*Post*Earn*	+/−	−0.13	0.03	−0.16
		(−3.58)***	(0.94)	(−3.31)***
Number of Obs.		4325	3716	
Adjusted R^2		0.10	0.06	
Panel B: *Cash Flows*$_{q+2}$				
Intercept	+/−	0.001	0.002	0.001
		(0.23)	(1.15)	(1.07)
Post	+/−	−0.002	−0.006	0.004
		(−0.62)	(−1.91)**	(0.87)
Earn	+/−	0.38	0.21	0.17
		(16.51)***	(12.09)***	(5.64)***
*Post*Earn*	+/−	−0.07	0.04	−0.11
		(−2.24)**	(1.46)	(−2.62)***
Number of Obs.		4110	3531	
Adjusted R^2		0.10	0.07	
Panel C: *Cash Flows*$_{q+3}$				
Intercept	+/−	0.000	0.001	0.001
		(0.24)	(0.82)	(0.73)
Post	+/−	−0.002	−0.004	0.002
		(−0.57)	(−1.44)	(0.53)
Earn	+/−	0.29	0.17	0.12
		(13.89)***	(11.02)***	(4.77)***
*Post*Earn*	+/−	−0.04	0.03	−0.07
		(−1.24)	(1.12)	(−1.64)*
Number of Obs.		3900	3347	
Adjusted R^2		0.10	0.06	

(*continued on next page*)

TABLE 5 (continued)

Variable	Predicted Sign	SAB 101 Firms Coefficients (t-statistics)	Unaffected Firms Coefficients (t-statistics)	Difference in Coefficients (F-test)
Panel D: *Cash Flows*$_{q+4}$				
Intercept	+/−	0.000	0.00	0.00
		(0.10)	(0.51)	(0.43)
Post	+/−	−0.000	−0.002	0.002
		(−0.05)	(−0.91)	(0.55)
Earn	+/−	0.27	0.13	0.14
		(13.21)***	(9.12)***	(5.79)***
*Post*Earn*	+/−	0.07	0.02	0.03
		(2.05)**	(0.99)	(0.86)
Number of Obs.		3708	3165	
Adjusted R^2		0.09	0.04	

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

Variable Definitions:

$LeadCFO_{(i,q+t)}$ = accumulated cash flow from operations (quarterly Compustat data item 108) scaled by end-of-period assets (quarterly Compustat data item 44) for firm *i* in one, two, three, and four subsequent quarters;

Post = indicator variable equal to 1 if the firm quarter observation occurs during the post-adoption period, and 0 otherwise;

$Earn_{(i,q)}$ = earnings, measured as the net income (quarterly Compustat data item 69) for firm *i* in quarter *q*, scaled by end-of-period assets (quarterly Compustat data item 44); and

$Post{*}Earn_{(i,q)}$ = *Post* multiplied by *Earn*.

of revenue recognition imposed by SAB No. 101 reduced SAB 101 firms' earnings informativeness, but did not impact unaffected firms. These findings are consistent with the FASB's belief that unearned revenues are informative.

Returns Analysis

Table 6 provides the results of estimating model (5), our regression of earnings announcement period returns on unexpected earnings. These ERC test findings corroborate the results from the cash flow analysis. For SAB 101 firms, we find a positive and statistically significant coefficient on *ΔEarn*, indicating that the market reacts to unexpected earnings information. We also find a negative and statistically significant coefficient on *Post*ΔEarn*, indicating that the informativeness of unexpected earnings declines during the post-adoption period.

For unaffected firms, we also find a significant positive coefficient on *ΔEarn*, indicating that unexpected earnings are associated with returns for these firms. However, the coefficient on *Post*ΔEarn* is not statistically significant, indicating no discernable change in the association between unexpected earnings and returns in the post-period.

The F-tests, reported in column 3, indicate that in the pre-adoption period SAB 101 firms had more informative unexpected earnings than unaffected firms. In addition, we find that SAB No. 101 adoption is associated with a significantly larger reduction in unexpected earnings informativeness for SAB 101 firms than for the unaffected firms.

P28980

TABLE 6
Coefficients (t-Statistics) from a Firm and Quarterly Fixed-Effects Regression of Earnings Announcement Period Returns on Unexpected Earnings for a Sample of Firms Affected by SAB No. 101 and a Sample of Unaffected Firms

$$Returns_{(i,t)} = \alpha + \beta_1 Post + \beta_2 \Delta Earn_{(i,q)} + \beta_3 Post^{*}\Delta Earn_{(i,q)} + \varepsilon$$

Variable	Predicted Sign	SAB 101 Firms Coefficients (t-statistics)	Unaffected Firms Coefficients (t-statistics)	Difference in Coefficients (F-test)
Intercept	+/−	−0.002	−0.000	−0.001
		(−0.78)	(−0.12)	(−0.44)
Post	+/−	−0.000	−0.000	0.002
		(−0.07)	(−0.13)	(0.07)
ΔEarn	+/−	0.02	0.01	0.011
		(5.61)***	(1.98)**	(1.94)**
*Post*ΔEarn*	+/−	−0.01	0.003	−0.014
		(−1.81)*	(0.62)	(−1.70)*
Number of Obs.		3091	2812	
Adjusted R^2		0.01	0.003	

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

Variable Definitions:

$Returns_{(i,t)}$ = two-day cumulative market-adjusted returns from CRSP (using the value weighted market index) for firm *i* on the earnings report date *t* and date $t-1$;

Post = indicator variable equal to 1 if the firm quarter observation occurs during the post-adoption period, and 0 otherwise;

$\Delta Earn_{(i,q)}$ = unexpected earnings, measured as the seasonally adjusted change in net income (quarterly Compustat data item 69) for firm *i* in period *q*, scaled by the end-of-period shares outstanding (quarterly Compustat data item 61); and

$Post^{*}\Delta Earn_{(i,q)}$ = *Post* multiplied by *ΔEarn*.

We conduct several different sensitivity analyses to ensure that our earnings informativeness results are not driven by our research design choices.[19] In our returns analysis, we increase the return window to five days and to ten days, and the results remain qualitatively the same (although the effect is weaker). We also scale unexpected earnings by assets (quarterly Compustat data item 44), define earnings as earnings excluding extraordinary items (quarterly Compustat data item 76), and as income from operations (quarterly Compustat data item 21) and the results remain unchanged.

Pro Forma Earnings

In Table 7, we report the results of models (6) and (7), our regressions of future annual cash flows and annual returns on pro forma earnings and accelerated earnings (the difference between previously reported and pro forma earnings). The first column of results indicates that both pro forma earnings and accelerated earnings are associated with future cash flows.

---

[19] To ensure that our results are not driven by outliers, we winsorize all variables in the reported regressions at five standard deviations. Truncating at the same level produces similar results. Similar results are also obtained when we use three standard deviations to winsorize the data. The results without winsorizing are similar to those reported.

TABLE 7
Coefficients and (t-Statistics) from Regressions of Annual One-Year-Ahead Cash from Operations and of Annual Returns on Pro Forma Earnings and Accelerated Earnings for SAB 101 Firms

$$Annual\ LeadCFO_{(i,t+1)} = \alpha + \beta_1 Pro\ Forma\ Earn_{(i,t)} + \beta_2 Accelerated\ Earn_{(i,t)} + \varepsilon$$

$$Annual\ Returns_{(i,t)} = \alpha + \beta_1 Pro\ Forma\ Earn_{(i,t)} + \beta_2 Accelerated\ Earn_{(i,t)} + \varepsilon$$

Variable	Predicted Sign	*Annual LeadCFO*	*Annual Returns*
Intercept	+/−	0.004 (0.24)	0.36 (3.56)
Accelerated Earn	+/−	0.41 (1.88)*	1.63 (1.70)*
Pro Forma Earn	+/−	0.58 (11.07)***	0.57 (1.67)*
Number of Obs.[a]		184	151
Adjusted R^2		0.396	0.007

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

[a] Of the 110 firms that provided pro forma data, approximately 80 percent provided pro forma earnings for all periods affected by the accounting change, and 70 percent had one-year ahead returns data available for each pro forma period presented. Consequently, our pro forma tests use 184 observations for the cash flow analysis and 151 for the returns analysis.

Variable Definitions:

Annual LeadCFO$_{(i,t+1)}$ = annual cash flow from operations (annual Compustat data item 308) scaled by end-of-period assets (annual Compustat data item 6) for firm *i* in year *t*+1;

Annual Returns$_{(i,t)}$ = annual cumulative market-adjusted returns from CRSP (using the value-weighted market return) for firm *i* from end of month four of fiscal year *t* until end-of-month four of fiscal year *t*+1;

Pro Forma Earn$_{(i,t)}$ = difference between actual earnings reported for firm *i* in period *t* and the amount of the pro forma adjustment associated with a cumulative effect from an accounting change related to SAB No. 101, scaled by end-of-period assets; and

Accelerated Earn$_{(i,t)}$ = amount of earnings related to periods after fiscal year 2000 that had been recognized in a prior period earnings (i.e., earnings that had been accelerated in periods prior to fiscal year 2000) scaled by end-of-period assets.

Furthermore, there is no statistically significant difference between the coefficients on these two variables (not tabulated), suggesting that the accelerated revenue targeted by SAB No. 101 and other revenue reported by these firms have similar associations with future cash flows. These results are consistent with the cash flow results reported in Table 5.

In the second column of Table 7, we find a significant association between pro forma earnings and annual returns, and accelerated earnings and annual returns. However, we find no statistically significant difference between the coefficients on accelerated versus pro forma earnings.[20]

Overall, the results of the cash flow, ERC, and pro forma tests provide evidence that SAB No. 101 adoption is associated with a reduction in the informativeness of earnings. These results are consistent with the FASB's argument that revenues recognized prior to completion of the earnings process provide relevant information about future performance resulting in more informative earnings.

---

[20] We use a long return window that begins on the first day of the fifth month of fiscal year *t* and ends at the end of the fourth month of fiscal year *t*+1 in order to incorporate the market's response to earnings on the earnings announcement date, on the date of the press conference discussing earnings, or on the 10-K filing date. When we begin our accumulation period at the end of the third month of period *t* and finish at the end of the third month of period *t*+1, our results are quantitatively the same.

Supplementary Analysis

The results of our earnings management tests suggest that some of the SAB 101 firms chose to accelerate revenue recognition to manage earnings. Despite this finding, the results of our earnings informativeness tests suggest that, on average, accelerated revenues are value relevant. However, this average result is not necessarily applicable for firms that accelerate revenue recognition to manage earnings. That is, firms could be managing earnings either to "mask true performance" or to provide value-relevant information to investors and thus make earnings more informative.

To provide evidence on these two alternatives, we conduct a supplemental analysis that uses the predicted probabilities from our logistic regression to partition our SAB 101 firms into those with higher versus those with lower incentives to accelerate revenue recognition to manage earnings.[21] We interact this partitioning variable with the earnings variable in the cash flow and ERC models to try to determine whether the extent of firms' earnings management incentives affects the informativeness of their earnings. We employ the following regressions for the pooled sample, controlling for fixed firm and quarter effects:

$$\begin{aligned} LeadCFO_{(i,q-1)} = {} & \alpha + \beta_1 Post + \beta_2 SAB + \beta_3 Post{*}SAB + \beta_4 High{*}SAB \\ & + \beta_5 Low{*}SAB + \beta_6 High{*}SAB{*}Post + \beta_7 Low{*}SAB{*}Post \\ & + \beta_8 Earn_{(i,q)} + \beta_9 Post{*}Earn_{(i,q)} + \beta_{10} Noincent{*}Earn \\ & + \beta_{11} Noincent{*}Post{*}Earn + \beta_{12} High{*}Earn \\ & + \beta_{13} High{*}Post{*}Earn + \beta_{14} Low{*}Earn + \beta_{15} Low{*}Post{*}Earn + \varepsilon \end{aligned} \tag{8}$$

$$\begin{aligned} Returns_{(i,t)} = {} & \alpha + \beta_1 Post + \beta_2 SAB + \beta_3 Post{*}SAB + \beta_4 High{*}SAB + \beta_5 Low{*}SAB \\ & + \beta_6 High{*}SAB{*}Post + \beta_7 Low{*}SAB{*}Post + \beta_8 \Delta Earn_{(i,q)} \\ & + \beta_9 Post{*}\Delta Earn_{(i,q)} + \beta_{10} Noincent{*}\Delta Earn_{(i,q)} \\ & + \beta_{11} Noincent{*}Post{*}\Delta Earn_{(i,q)} + \beta_{12} High{*}\Delta Earn_{(i,q)} \\ & + \beta_{13} High{*}Post{*}\Delta Earn_{(i,q)} + \beta_{14} Low{*}\Delta Earn_{(i,q)} \\ & + \beta_{15} Low{*}Post{*}\Delta Earn_{(i,q)} + \varepsilon \end{aligned} \tag{9}$$

where:

$LeadCFO_{(i,q+1)}$ = cash flow from operations (quarterly Compustat data item 108) scaled by assets (quarterly Compustat data item 44) for firm i in quarter $q+1$;

$Returns_{(i,t)}$ = two-day cumulative market adjusted returns from CRSP (using the value weighted index return as our measure of the market's return) for firm i on the earnings report date t and date $t-1$;

Post = indicator variable equal to 1 if the firm quarter observation occurs during the post-adoption period, and 0 otherwise;

---

[21] Since there are firms in our sample that are missing data necessary for the logistic analysis, we code these firms as having unknown incentives to manage earnings and create a dummy variable (*Noincent*) to measure the informativeness of earnings for these firms.

SAB = indicator variable equal to 1 for SAB No. 101 firms, and 0 otherwise;
$High$ = indicator variable equal to 1 for SAB No. 101 firms with above the median-fitted value from the logistic regression of earnings management incentives, and 0 otherwise;
Low = indicator variable equal to 1 for SAB No. 101 firms with below the median-fitted value from the logistic regression of earnings management incentives, and 0 otherwise;
$Noincent$ = indicator variable equal to 1 for SAB No. 101 firms missing data necessary to estimate logistic regression of earnings management incentives, and 0 otherwise;
$Earn_{(i,q)}$ = earnings, measured as the net income (quarterly Compustat data item 69) for firm i in quarter q, scaled by the end-of-period assets (quarterly Compustat data item 44); and
$\Delta Earn_{(i,q)}$ = Unexpected earnings, measured as the seasonally adjusted change in net income (quarterly Compustat data item 69) for firm i in quarter q, scaled by the end-of-period shares outstanding (quarterly Compustat data item 61).
$*$ = Indicates multiplication.

The results of this supplemental analysis are reported in Table 8. Panel A provides associations between earnings and one-quarter-ahead cash flows. The statistically significant coefficients on the interaction of *Earn* with each earnings management indicator variable (*High*, *Low*, and *Noincent*) indicate a positive association between earnings and one-quarter-ahead cash flows in the pre-adoption period. This suggests that regardless of firms' earnings management incentives, there is a higher association between earnings and future cash flows when firms accelerate revenue recognition. Our results also indicate that, on average, in the post-adoption period SAB 101 firms with both high and low earnings management incentives experience a greater reduction in the ability of earnings to predict future cash flows than do unaffected firms.

These results are consistent with Subramanyan's (1996) argument that some firms manage earnings to provide value-relevant information to shareholders. They also suggest that the FASB's argument that accelerating the recognition of unearned revenue increases the value relevance of earnings holds for firms that have high incentives to manage earnings.

Panel B of Table 8 provides the results of our supplemental analysis when we use returns as the metric for earnings informativeness. We find that SAB 101 firms with high earnings management incentives have more informative unexpected earnings than do unaffected firms. These high-incentive SAB 101 firms also experience a significant deterioration in the association between unexpected earnings and announcement period returns in the post-adoption period. Like the cash flow tests, these results are consistent with Subramanyan's (1996) signaling hypothesis. However, in contrast to the cash flow tests, we do not find similar results for the low-incentive firms. Thus, even though our supplemental cash-flow tests suggest that earnings of low-incentive firms are more informative than the earnings of unaffected firms, the market does not appear to have a stronger reaction. It is difficult to draw any conclusions from the mixed results on the low-incentive firms.

Taken as a whole, the results from these supplemental analyses provide limited evidence that the earnings numbers reported by SAB 101 firms prior to SAB No. 101 adoption were more informative than earnings in the post-period, regardless of the earnings management incentives.

TABLE 8

Coefficients and (t-statistics) from a Firm and Quarterly Fixed-Effects Regression of One-Quarter-Ahead Cash Flows on Earnings for a Sample of Unaffected Firms and a Sample of SAB 101 Firms Partitioned on Earnings Management Incentives

Panel A: CFO Analysis

$$
\begin{aligned}
LeadCFO_{(i,q+1)} = {} & \alpha + \beta_1 Post + \beta_2 SAB + \beta_3 Post{*}SAB + \beta_4 High{*}SAB + \beta_5 Low{*}SAB \\
& + \beta_6 High{*}SAB{*}Post + \beta_7 Low{*}SAB{*}Post + \beta_8 Earn_{(i,q)} + \beta_9 Post{*}Earn_{(i,q)} \\
& + \beta_{10} Noincent{*}Earn + \beta_{11} Noincent{*}Post{*}Earn + \beta_{12} High{*}Earn \\
& + \beta_{13} High{*}Post{*}Earn + \beta_{14} Low{*}Earn + \beta_{15} Low{*}Post{*}Earn + \varepsilon
\end{aligned}
$$

Variable	Predicted Sign	Coefficient/(t-stat)
Intercept	+/−	−0.010 (−1.97)
Post	+/−	−0.009 (−1.43)
SAB	+/−	−0.004 (−0.50)
*Post*SAB*	+/−	0.015 (1.00)
*High*SAB*	+/−	−0.012 (−2.14)**
*Low*SAB*	+/−	0.012 (1.46)
*High*SAB*Post*	+/−	0.036 (3.61)***
*Low*SAB*Post*	+/−	−0.033 (−2.34)***
Earn	+	0.27 (8.10)***
*Post*Earn*	+/−	0.019 (0.34)
*Noincent*Earn*	+/−	0.249 (2.52)***
*Noincent*Post*Earn*	+/−	0.132 (0.66)
*High*Earn*	+/−	0.299 (5.11)***
*High*Post*Earn*	+/−	−0.236 (−2.76)***
*Low*Earn*	+/−	0.451 (5.15)***
*Low*Post*Earn*	+/−	−0.225 (−1.75)**

(*continued on next page*)

TABLE 8 (continued)

Panel B: Returns Analysis

$$\begin{aligned} Returns_{(i,t)} = {} & \alpha + \beta_1 Post + \beta_2 SAB + \beta_3 Post{*}SAB + \beta_4 High{*}SAB + \beta_5 Low{*}SAB \\ & + \beta_6 High{*}SAB{*}Post + \beta_7 Low{*}SAB{*}Post + \beta_8 \Delta Earn_{(i,q)} + \beta_9 Post{*}\Delta Earn_{(i,q)} \\ & + \beta_{10} Noincent{*}\Delta Earn_{(i,q)} + \beta_{11} Noincent{*}Post{*}\Delta Earn_{(i,q)} + \beta_{12} High{*}\Delta Earn_{(i,q)} \\ & + \beta_{13} High{*}Post{*}\Delta Earn_{(i,q)} + \beta_{14} Low{*}\Delta Earn_{(i,q)} + \beta_{15} Low{*}Post{*}\Delta Earn_{(i,q)} + \varepsilon \end{aligned}$$

Variable	Predicted Sign	Coefficient/(t-stat)
Intercept	+/−	−0.001 (−0.41)
Post	+/−	−0.00 (−0.14)
SAB	+/−	0.002 (0.33)
*Post*SAB*	+/−	−0.009 (−0.97)
*High*SAB*	+/−	−0.002 (−0.68)
*Low*SAB*	+/−	0.003 (0.58)
*High*SAB*Post*	+/−	0.003 (0.70)
*Low*SAB*Post*	+/−	−0.004 (−0.57)
ΔEarn	+	0.009 (1.95)**
*Post*ΔEarn*	+/−	0.003 (0.63)
*Noincent*ΔEarn*	+/−	0.015 (1.70)**
*Noincent*Post*ΔEarn*	+/−	−0.032 (−1.79)**
*High*ΔEarn*	+/−	0.01 (2.17)**
*High*Post*ΔEarn*	+/−	−0.021 (−2.19)**
*Low*ΔEarn*	+/−	−0.002 (−0.26)
*Low*Post*ΔEarn*	+/−	0.01 (0.91)

***,**,* Indicates significance at the 1 percent, 5 percent, and 10 percent levels, respectively, based on a one- or two-tailed test as appropriate.

Variable Definitions:

Returns$_{(i,t)}$ = two-day cumulative market adjusted returns from CRSP (using the value-weighted index return as our measure of the market's return) for firm *i* on the earnings report date *t* and date *t*−1;

LeadCFO$_{(i,q+1)}$ = cash flow from operations (quarterly Compustat data item 108) scaled by assets (quarterly Compustat data item 44) for firm *i* in quarter *q*+1;

Post = indicator variable equal to 1 if the firm quarter observation occurs during the post-adoption period, and 0 otherwise;

SAB = indicator variable equal to 1 for SAB 101 firms, and 0 otherwise;

(*continued on next page*)

TABLE 8 (continued)

High = indicator variable equal to 1 for SAB 101 firms with above the median fitted value from the logistic regression of earnings management incentives, and 0 otherwise;
Low = indicator variable equal to 1 for SAB 101 firms with below the median fitted value from the logistic regression of earnings management incentives, and 0 otherwise;
Noincent = indicator variable equal to 1 for SAB 101 firms missing data necessary to estimate logistic regression of earnings management incentives, and 0 otherwise;
$Earn_{(i,q)}$ = earnings, measured as the net income (quarterly Compustat data item 69) for firm i in quarter q, scaled by the end-of-period assets (quarterly Compustat data item 44);
$\Delta Earn_{(i,q)}$ = unexpected earnings, measured as the seasonally adjusted change in net income (quarterly Compustat data item 69) for firm i in quarter q, scaled by the end-of-period shares outstanding (quarterly Compustat data item 61); and
* = indicates multiplication.

VII. CONCLUSIONS

In this study we examine a sample of firms that were required to adjust their earnings because of SAB No. 101, an accounting bulletin issued to clarify the SEC's interpretation of revenue recognition rules. Specifically, we compare the pre-adoption earnings properties of a sample of SAB 101 companies to their post-adoption earnings properties, and we compare the SAB 101 firms to unaffected firms that are in the same industry, have similar asset sizes, and engage in the same types of revenue generating activities as the SAB 101 firms.

The SEC argues that improper revenue recognition is the single largest cause of financial reporting problems and that firms use premature revenue recognition to meet Wall Street expectations. Based on the SEC's argument, we consider whether firms that accelerated revenue recognition were more likely to meet earnings benchmarks and had greater earnings management incentives.

Critics of SAB No. 101 argue that the bulletin resulted in changes in GAAP that made the financial reports of the firms affected by SAB No. 101 less informative. The FASB echoed this concern in justifying its ongoing revenue recognition project by arguing that revenues recorded prior to the completion of the earnings process contain value-relevant information about future performance. We examine this possibility by comparing the associations between earnings and cash flows and unexpected earnings and announcement period returns for SAB 101 and unaffected firms in the periods prior and subsequent to SAB No. 101 adoption.

We find evidence that firms recording SAB No. 101 adjustments were more likely to meet important earnings benchmarks in the pre-adoption period than in the post-adoption period, and that SAB 101 firms were more likely than unaffected firms to meet earnings benchmarks. However, we also find that the association between earnings and next period cash flows is higher for SAB 101 firms than for unaffected firms and that the association between earnings and future cash flows declines for SAB 101 firms in the post-adoption period. We find similar results in our analysis of the relationship between unexpected earnings and announcement period returns.

We also find that, when we partition firms that accelerate revenue recognition into groups that have strong and weak incentives to manage earnings, the earnings of firms with the strongest incentive to manage earnings have more informative earnings than firms unaffected by the regulation. This result is consistent with the FASB's assertion that the relevance of reported earnings is, on average, greater when revenue recognition is not delayed until the completion of the earnings process, even when earnings management incentives are high.

The results of this study provide information of particular interest to practitioners, regulators, and standard setters who, when developing and implementing revenue recognition standards, must consider the trade-offs between discretion in reporting revenues prior to the completion of the earnings process, and the informativeness of unearned revenues.

APPENDIX

Panel A: License Agreement—Millennium Pharmaceutical

Revenue recognition policy prior to SAB No. 101:

"The Company recognized licensing revenue when underlying performance criteria for payment had been met and when the Company had an unconditional right to such payment. Depending on a license agreement's terms, recognition criteria may have been satisfied upon achievement of milestones, passage of time or product sales by licensee."

Revenue recognition policy after SAB No. 101 adoption:

"Under SAB No. 101, the Company recognizes revenue from licensing agreements using a contingency-adjusted performance model. Under this method, revenue related to up-front, time-based, and performance-based licensing payments are recognized over the entire contract performance period ... Payments received by the Company in excess of amounts earned are classified as deferred revenue."

Financial Statement Effects of Adjustment (000s)

	1998	1999	2000
Earnings	10,338	(379,904)	(201,927)
Cumulative Effect			(107,692)
Earnings after Cumulative Effect			(309,619)
Pro forma Earnings	(10,461)	(417,147)	
Amount of earnings accelerated	20,799	37,243	

Panel B: Installment and Customer Acceptance—Integrated Measurement Systems

Change in Revenue Recognition Policy:

"In the fourth quarter of 2000, we implemented SAB No. 101, effective January 1, 2000, by changing the timing of revenue recognition for system sales from the historical industry practice of when product is shipped (title passes) to when a customer has accepted a shipped product. This change in our revenue recognition policy has been accounted for as a change in accounting principle effective January 1, 2000. We have recorded an after-tax non-operating charge against net income in the amount of $7.3 million reflecting the deferral of revenue net of related expenses for shipments of our products previously reported as revenue prior to January 1, 2000 that had not been accepted by customers as of December 31, 1999. Billings in advance of revenue recognition are recorded as deferred revenues and the related costs are recorded as inventory at customer sites and as deferred commissions."

Financial Statement Effects of Adjustment (000s)

	1998	1999	2000
Earnings	(3,331)	5,851	10,317
Cumulative Effect			(7,317)
Earnings after Cumulative Effect			3,000
Pro forma Earnings	(3,850)	3,033	
Amount of earnings accelerated	519	2,818	

REFERENCES

Ameen, P. D. 2000. Letter from Committee on Corporate Reporting of the Financial Executives Institute. Available at: http://www.fei.org/download/SAB_Delay_ltr.pdf.

Barth, M., D. Cram, and K. Nelsen. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (January): 27–58.

Beaver, W. H., M. McNichols, and K. Nelson. 2003. An alternative interpretation of the discontinuity in earnings distributions. Working paper, Stanford University.

Begley, J., and R. Freedman. 2004. The changing role of accounting numbers in public lending agreements. *Accounting Horizons* 18 (December): 81–107.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (December): 99–126.

———. 2004. Commentary on an alternative interpretation of the discontinuity in earnings distributions. Working paper, University of Washington.

Bushmen, R., Q. Chen, E. Engel, and A. Smith. 2002. Financial accounting information, organizational complexity and corporate governance systems. Working paper, University of Chicago.

Dechow, P., S. P. Kothari, and R. Watts. 1998. The relation between earnings and cash flow. *Journal of Accounting and Economics* 25 (May): 131–168.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72: 1–33.

Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprises.* Statement of Financial Accounting Concepts No. 1. Norwalk, CT: FASB.

———. 1985. *Elements of Financial Statements.* Statement of Financial Accounting Concepts No. 6. Norwalk, CT: FASB.

Graham, J., C. Harvey, and S. Rajgopal. 2004. The economic implications of corporate financial reporting. Working paper, Duke University.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31 (September): 105–150.

Lev, B. 1989. On the usefulness of earnings and earnings research: Lessons and directions from two decades of empirical research. *Journal of Accounting Research* 27: 153–192.

Levitt, A. 1998. The numbers game. Speech delivered at the NYU Center for Law and Business, New York, NY, September 28. Available at: http://www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

Osterland, A. 2000. Hard lessons. In its ongoing war on earnings management the SEC may have outdone itself with new rules on revenue recognition. *CFO.com.*

PricewaterhouseCoopers. 2001. Revenue Is revenue, right? Under SAB No. 101, the answer is not always apparent. *Deal Flash* 11.

Rajgopal, S., T. Shevlin, and M. Venkatachalam. 2003. Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. Working paper, Duke University.

Securities and Exchange Commission (SEC). 1999. *Revenue Recognition.* Staff Accounting Bulletin No. 101. Washington, D.C.: Government Printing Office.

Subramanyan, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (May): 249–281.

Teoh, S., I. Welch, and T. Wong. 1998a. Earnings management and the post-issue performance of seasoned equity offerings. *Journal of Financial Economics* 50 (October): 63–99.

———, ———, ———. 1998b. Earnings management and the long-term market performance of initial public offerings. *Journal of Financial Economics* 53 (January): 1935–1974.

Turner, L., J. R. Dietrich, K. Anderson, and A. Bailey. 2001. Accounting restatements. Working paper, Colorado State University.

Watts, R. L., and J. L. Zimmerman. 1986. *Positive Accounting Theory.* Englewood Cliffs, NJ: Prentice Hall.

———, and ———. 1990. *Positive Accounting Theory*: A ten-year perspective. *The Accounting Review* 65: 131–156.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 403–421

Evidence That Investors Trade on Private Event-Period Information around Earnings Announcements

Orie E. Barron
Pennsylvania State University

David G. Harris
Syracuse University

Mary Stanford
Texas Christian University

ABSTRACT: Holthausen and Verrecchia's (1990) and Kim and Verrecchia's (1997) theoretical models predict that private information inferred at the time of an earnings announcement (private event-period information) is associated with greater trading volume. We provide empirical evidence consistent with these theories. Specifically, announcements that increase analysts' private information (as measured by Barron et al.'s [1998] empirical proxies) are associated with increased trading volume, consistent with some investors similarly acquiring private event-period information. In addition, announcements that decrease analysts' consensus are associated with more trading volume. Because consensus declines when private information increases, this finding provides reinforcing evidence that investors trade following earnings announcements because of private information that becomes useful only in conjunction with the information in the announcement and that this information is important enough to spur trading.

Keywords: *trading volume; private information; analysts' forecasts.*

Data Availability: *The data used in this study are commercially available from the sources described in the text.*

For their helpful comments, we thank Linda Bamber, Donal Byard, Sandy Callaghan, Dan Givoly, Henok Louis, Ray Pfeiffer, Pamela Stuerke, Richard Schneible, Charlie Smith, Ro Verrecchia, Bill Wempe, two anonymous reviewers, and workshop participants at the Syracuse University and the 2000 American Accounting Association Annual Meeting. We also gratefully acknowledge the contribution of I/B/E/S International Inc. for providing earnings per share forecast data, available through the Institutional Brokers' Estimate System. These data have been provided as part of a broad academic program to encourage earnings expectation research.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted April 2000
Accepted September 2004

I. INTRODUCTION

Trading volume is the most observable indicator that financial accounting information affects investors (Cready and Hurtt 2002). This article reports empirical evidence concerning the descriptive validity of two related theoretical models of such trading volume: Kim and Verrecchia (1997) (hereafter KV) and Holthausen and Verrecchia (1990) (hereafter HV). These models predict that public announcements, e.g. earnings announcements, may lead investors to disagree about firm value and that this disagreement is sufficiently material to stimulate trading. We provide empirical evidence consistent with KV's prediction that earnings announcements stimulate the development of private event-period information, which spurs trading. In addition, our evidence is consistent with the related HV prediction that consensus-decreasing earnings announcements lead to trading volume that is not evident around other earnings announcements.

KV and HV are important advances in the theoretical literature because they move beyond early models that assumed public information simply *substitutes* for private information (e.g., Kim and Verrecchia 1991). For example, in KV's model, investors trade around public information releases as a result of three types of information: (1) private information gathered in anticipation of the announcement (pre-announcement information); (2) public information contained in the announcement itself (value-relevant information); and (3) private event-period information. This third type of information allows for a *complementary* relation between public and private information, in that public announcements spur investors to develop private event-period information that leads investors to act (i.e., trade). Prior research has shown that trading volume is associated with the first two types of information (Bamber et al. 1997; Beaver 1968), but to date there is no empirical evidence on whether the third type of information spurs trading.[1]

KV define private event-period information as new information arising from an interaction between information contained in the public announcement and private information gathered elsewhere that becomes useful only in conjunction with the announcement itself. Barron et al. (2002, 824–825) provide the following example of private event-period information:

> For example, consider two analysts. The first is a Certified Public Accountant who understands when and how specific accruals will reverse and affect future earnings. The second, a former political economist, correctly expects that back-orders from a Brazilian customer will soon decline and decrease earnings. Neither of these analysts can use his or her indirect knowledge in isolation. Yet both can combine their knowledge with reported earnings to produce "uniquely privately inferred information."

Intuitively, the information contained in the earnings announcement enables analysts and investors to capitalize on their private information to draw new inferences regarding the value of the firm. Although Barron et al. (2002) provide empirical evidence that earnings announcements lead *analysts* to develop private event-period information in a manner consistent with KV's theory, there is no empirical evidence that private event-period information is sufficiently important to lead *investors* to trade.

---

[1] Prior empirical research has examined other theoretical predictions about how disagreement causes trading volume. For example, Ziebart (1990) finds that as dispersion in analysts' forecasts increases trading volume following earnings announcements increases. But dispersion in analysts' forecasts reflects both uncertainty and consensus, and both increases in uncertainty and decreases in consensus could lead to trading (Dontoh and Ronen 1993). Thus, it is unclear from Ziebart (1990) whether the consensus or the uncertainty component of the change in dispersion (or both) is associated with trading following earnings announcements. Atiase and Bamber (1994) use the level of forecast dispersion as a measure of pre-disclosure information asymmetry but they do not investigate asymmetry in event-period information. Similarly, Bamber et al. (1997) find that earnings announcements leading to differential forecast revisions also lead to higher trading. But as they point out, and our evidence confirms, differential forecast revisions are likely attributable to investors having pre-disclosure information of differential precision. Whether trade following earnings announcements is associated with private event-period information remains unresolved.

In a precursor to KV, HV derives a trading volume effect termed the "consensus" effect, which is similar to the private event-period information effect in KV. Investor consensus is the commonality of investors' information as gauged by the correlation of their forecast errors. HV predicts that volume will increase as consensus decreases. HV is more rudimentary than KV in that HV is silent on the source of (or reason for) the decrease in consensus. However, both theories make a similar prediction if consensus decreases as private information increases.[2] In such a case, a decrease in consensus following an earnings announcement indirectly reflects the presence of private event-period information.

Drawing on the Barron, Kim, Lim, and Stevens (1998) (hereafter BKLS) empirical proxies for *analysts'* private information and consensus (described in more detail later in the article), this study provides empirical evidence supporting the KV and HV predictions that *investors* trade when earnings announcements trigger the development of private information and decrease consensus. The BKLS proxy for private event-period information relies on more stringent assumptions than the BKLS consensus measure.[3] Consequently, obtaining similar inferences with both measures provides reinforcing evidence that investors trade following earnings announcements based on private information that becomes useful only in conjunction with the information in the announcement. Additionally, if the HV and KV theories are, in fact, valid, then our evidence is consistent with analysts' information serving as a proxy for some sophisticated investors' information.

An important consequence of the existence of trade based on private event-period information is its potential to increase a firm's cost of capital. In related work using BKLS empirical measures, Botosan et al. (2004) find that the equity cost of capital increases with the precision of analysts' private information. Barron et al. (2002) find that earnings announcements trigger the release of a large percentage of this private information. As Kim and Verrecchia (1994, 1997) contend, it is trade based on this type of private information that adversely affects firms' costs of capital. Thus, our evidence contributes to our understanding of how earnings announcements affect both investors' trading decisions and, potentially, firms' cost of capital.

In sum, we show that some investors trade around earnings announcements in a manner consistent with their possessing private event-period information. The results support the descriptive validity of the KV/HV theoretical predictions, including the somewhat counterintuitive notion that earnings announcements can lead to *more* private information, which, in turn, *reduces* consensus about firm value and is sufficiently material to spur investors to trade. In addition to supporting theoretical predictions that are somewhat contrary to conventional wisdom, the results also have important implications for firm managers and for accounting policymakers seeking to level the informational playing field. Specifically, if a significant amount of trade on private information occurs when public disclosure peaks, then some unsophisticated traders are likely at a greater disadvantage. Consistent with Botosan et al. (2004), this uneven informational playing field may increase firms' cost of capital.

---

[2] Empirical tests reported later in this article confirm that our proxies for consensus and private event-period information are significantly negatively correlated. Botosan et al. (2004) also document a significant negative relation between the precision of analysts' private information and consensus.

[3] See Barron et al. (2002) for evidence concerning the reasonableness of the BKLS assumptions.

The paper proceeds as follows: Section II describes the two theoretical models of trading and develops the hypotheses. Section III specifies the empirical models and then describes the variable measurement. Section IV describes the sample and discusses the results. Section V provides sensitivity analyses that serve to link this study to prior empirical research. Section VI concludes the paper.

II. THEORETICAL BASES AND HYPOTHESES

Our study tests the predictions of two widely cited theories of how information affects trading volume: HV and KV. Because of differences in their assumptions, these two analytical models are not directly comparable; however, both models make similar predictions about how private event-period information will affect trading volume around public announcements (e.g., earnings announcements).

In KV, public announcements stimulate three kinds of information-related trading. First, some trading arises because investors have private pre-disclosure information of differential precision. Investors with more precise private information place more reliance on that information and less reliance on the earnings announcements. Thus, there is more (less) differential belief revision around earnings announcements with more (less) differential precision, and this differential belief revision stimulates more (less) trading (see also Kim and Verrecchia 1991).

Second, when the precision (amount) of investors' private pre-disclosure information differs, earnings announcements that contain *value-relevant information* stimulate trading. The KV model demonstrates that this portion of announcement period trading volume is positively related to contemporaneous price changes.

Third, some trading arises because of *private event-period information* investors develop in conjunction with the content of the earnings announcement itself. In other words, investors trade in the wake of earnings announcements not just because of the value-relevance of the information or differences in their prior beliefs, but also because their private event-period information leads them to interpret the reported amounts differently than others who lack this information. KV characterize this third type of trading volume as "residual trading," that is, information-induced trading not positively related to differences in the precision of investors' private pre-disclosure information (the first category of information-related trading) or to contemporaneous price changes (the second category of information-related trading). This third source of trading is similar in spirit to the consensus effect in HV, in that consensus decreases when investors rely more on private information and less on public information.

In the HV model, investors' interpretations of a public announcement contain both common and idiosyncratic error components so that heterogeneous interpretations of a public announcement lead to price and volume reactions. HV examine an effect of information on trading called the "consensus" effect. Consensus refers to the extent to which investors share common information. The more investors rely on common (public) information, the more their forecast errors exhibit covariance. Conversely, a decrease in the covariance of forecast errors indicates a decrease in reliance on public information. HV predicts that consensus-decreasing news events cause trading volume.

One limitation of HV is that it is a single-period model. Investors observe a signal about the value of the risky asset and then trade on the basis of their inferences. HV assumes that investors have homogeneous prior expectations and that the precision of their pre-disclosure information is homogeneous. Because HV assumes all investors have homogeneous prior expectations (i.e., prior consensus is equal to 1), public announcements cannot increase consensus across investors.[4]

In contrast, the subsequent KV (1997, 396) study introduces a two-period model they describe as "more descriptive of real market settings." The predictions in KV are consistent with HV, but because KV has a pre-disclosure period in which investors can disagree, consensus can increase or decrease as a result of an earnings announcement. Consistent with the KV assumptions, the descriptive evidence reported in Table 1 of this article indicates that consensus decreases following earnings announcements for approximately 60 percent of our sample. This evidence is consistent with KV's assertion that their model is more descriptive of actual earnings announcements.

The first hypothesis concerns the "residual trading" in KV, that is, trading triggered by the third category of information: private event-period information. The second hypothesis predicts a negative relation between trading volume and consensus. Evidence consistent with both hypotheses provides reinforcing empirical support for KV's theoretical prediction that part of the trading volume response to earnings announcements arises because the earnings announcement leads investors to develop private event-period information. The two hypotheses, stated in alternate form, follow:

H1: Earnings announcement "residual" trading volume that is unexplained by price changes or differential pre-disclosure private information is positively associated with private event-period information.

H2: Earnings announcement "residual" trading volume that is unexplained by price changes or differential pre-disclosure private information is negatively associated with changes in consensus.

III. RESEARCH DESIGN AND VARIABLE MEASUREMENT

This section describes the models and the empirical proxies we use to test the hypotheses. Drawing on the BKLS empirical proxies for private information and consensus we use *analysts'* (observable) belief revisions to draw inferences about *investors'* (unobservable) private event-period information. Specifically, we examine how *quarterly* earnings announcements affect properties of analysts' forecasts of *annual* earnings. In particular, we examine how the BKLS proxies for: (1) the precision (amount) of private information in analysts' forecasts and (2) analysts' consensus about annual earnings change following each quarterly earnings announcement. To isolate residual trading volume we control for trading associated with price changes and differential belief revisions arising from differences in the precision of pre-disclosure private information. Finding that increases in the precision of analysts' private information and decreases in analysts' consensus are associated with residual trading volume would support the theoretical proposition that part of the trading volume around earnings announcements arises because *investors* trade based on private event-period information.

We use the following empirical model to test H1:

---

[4] Conversely, in models in which event-period information is entirely common, such as Kim and Verrecchia (1991), public announcements cannot decrease consensus. Such models are not descriptive of the U.S. market given evidence in Barron et al. (2002) that earnings announcments typically decrease the BKLS measure of consensus.

$$EXVOL_j = a_0 + a_1 \Delta PPRECISION_j + a_2 LARETURN_j + a_3 LJUMBLING_j + e_j \quad (1)$$

where:

$EXVOL_j$ = event-period-induced excess trading volume, defined as the natural log of the cumulative trading volume over the three-day event window (days −1 to +1 around the interim earnings announcement) minus the natural log of the firm-specific median trading volume, where median volume is the median volume for contiguous three-day periods over the 249 trading days prior to the earnings announcement event window, where j is the firm-quarter subscript;

$\Delta PPRECISION_j$ = a continuous variable measuring changes in the BKLS proxy (defined below) for the precision of analysts' private information;

$LARETURN_j$ = the natural log of the absolute value of the sum of firm's daily returns in the three-day event window;

$LJUMBLING_j$ = the natural log of 1 minus the Pearson correlation between: (1) the annual earnings forecasts issued by individual analysts in the 45 days before the interim earnings announcement; and (2) the annual earnings forecasts issued by the same analysts within the 30 days after the interim earnings announcement; and

e_j = the error term for firm-quarter j.

In HV, trading occurs only when consensus decreases. Thus, we test H2 with an indicator variable set to 1 when consensus declines as well as a continuous variable. We use the following models to test H2:

$$EXVOL_j = b_0 + b_1 CONNEG_j + b_2 LARETURN_j + b_3 LJUMBLING_j + e_j \quad (2)$$

$$EXVOL_j = c_0 + c_1 \Delta CONSENSUS_j + c_2 LARETURN_j + c_3 LJUMBLING_j + e_j \quad (3)$$

where all variables are as defined above, and:

$CONNEG_j$ = an indicator variable set to 1 if the BKLS proxy (defined below) for analysts' consensus declined around the quarterly earnings announcement, and 0 otherwise; and

$\Delta CONSENSUS_j$ = a continuous variable measuring changes in the BKLS proxy (defined below) for analysts' consensus around the quarterly earnings announcement.

The KV model characterizes residual trading volume as that portion of announcement-period volume not caused by (1) value-relevant information reflected in contemporaneous price changes, (2) differences in the precision of investors' pre-disclosure private information reflected in differential belief revisions, and (3) non-informational (e.g., liquidity) factors. Accordingly, our tests control for trading volume arising from these other factors. We control for first factor, value-relevant information, with a measure of the absolute return (*LARETURN*) around the earnings announcement. We control for the second factor, differential prior precision of investors' information, with a measure of belief jumbling

(*LJUMBLING*). We control for the third factor, non-information-based trading, by subtracting firm-specific median trading volume from the dependent variable, *EXVOL*. This methodology is similar to the approach Atiase and Bamber (1994) and Kross et al. (1994) use to abstract from non-announcement, firm-specific levels of liquidity trading and ongoing levels of information-based trading unrelated to the theoretical notion of announcement-period residual volume.[5,6]

The explanatory variables of interest, $\Delta PPRECISION$, *CONNEG*, and $\Delta CONSENSUS$, are based on the BKLS proxies detailed in the next section; this is followed by a discussion of the control variables. Due to extreme values in the empirical proxy *PPRECISION*, we estimate Equation (1) as a rank regression, with each variable being ranked within each quarter, while Equations (2) and (3) are estimated with ordinary least squares.

Empirical Proxies for Private Event-Period Information and Consensus

This section describes our empirical proxies for KV's theoretical construct, the precision, accuracy, or amount of private information, which we label *PPRECISION*, and HV's theoretical consensus construct or the commonality of investors' information, which we label *CONSENSUS*.

KV define the amount of private event-period information as the precision of investors' private information signals where precision is the inverse of the variance of the signal around its realization (or signal accuracy). Note that as analysts possess more information, precision increases. HV define investor consensus as the commonality of investors' information, gauged by the correlation of their forecast errors around an assumed expected error of zero (i.e., expectations are assumed to be unbiased).

We empirically measure the extent of private event-period information and consensus with the following BKLS proxies:

$$PPRECISION = \frac{D}{V^2} = (1 - CONSENSUS) \cdot \left(\frac{1}{V}\right) \tag{4}$$

$$CONSENSUS = 1 - D/V \tag{5}$$

where:

D = dispersion in analysts' forecasts i.e., the sample variance of the individual forecasts (FC_i) around the mean forecast ($\overline{FC}$), measured as $\sum_{i=1}^{n} (FC_i - \overline{FC})^2/(n - 1)$, where n is the number of forecasts; and

V = the mean of the squared differences between individual analysts' forecasts (FC_i) and reported earnings per share (EPS) measured as $\sum_{i=1}^{n} (FC_i - EPS)^2/n$.

---

[5] Ross (1989) and Kandel and Pearson (1995) considered it unlikely that liquidity needs would be large enough to account for observed ongoing levels of security trading. Abstracting from the average non-announcement level of trading will likely abstract some *ongoing* information-based trading. Thus, our measure is a conservative estimate of information-based trading around earnings announcements.

[6] Our findings are qualitatively identical when we use the natural log of the cumulative trading volume over the three-day event window as the dependent variable and the natural log of the median preannouncement volume as an additional control variable.

To understand the intuition behind these BKLS measures assume the number of forecasts (n) is large and consider the extreme cases of D/V equal to 0 or 1. First, note that with a large number of forecasts, error due to common information is simply the difference between the mean forecast ($\overline{FC}$) and realized earnings per share (EPS), i.e., idiosyncratic error is completely averaged out of the mean leaving only common error. Second, error due to private information is the difference between individual forecasts (FC_i) and the mean forecast ($\overline{FC}$). $D/V = 1$ and *CONSENSUS* equals 0 when the mean forecast equals realized earnings. In this case, all forecast error is idiosyncratic and the BKLS model suggests that forecasts are based entirely on private information. In contrast, $D/V = 0$ and *CONSENSUS* equals 1 when all forecasts are exactly equal to the mean forecast. In this case, all forecast error is common and the BKLS model suggests that forecasts are based entirely on common information. In the unlikely case that all forecasts exactly equal realized earnings consensus is undefined because V is equal to 0 and D/V is infinite, consistent with perfectly accurate information.

A close examination of Equations (4) and (5) reveals that the precision (or accuracy) of private information (*PPRECISION*) and consensus (*CONSENSUS*) are closely related but not identical constructs. Specifically, *CONSENSUS* measures the degree of common information relative to total information. *CONSENSUS* can decrease either due to an *increase* in the amount of private information or a *decrease* in the amount of common information. In contrast, *PPRECISION* reflects the absolute precision (or accuracy) of information ($1/V$) multiplied by the relative degree to which this information is private in nature (D/V). *PPRECISION* can only increase because the accuracy of private information increased.[7] Consistent with the HV and KV theoretical constructs representing related but different concepts, these two measures, *PPRECISION* and *CONSENSUS*, reflect private information in different ways.[8]

In Regression (1) Δ*PPRECISION* is the difference between *PPRECISION* based on annual earnings forecasts issued by individual analysts in the 45 days before the interim earnings announcement and *PPRECISION* estimated with annual earnings forecasts issued by the same analysts within the 30 days after the interim earnings announcement. As mentioned above, because *PPRECISION* is unbounded, Regression (1) is estimated using rank regression with each variable ranked within each quarter.

In Regression (2) *CONNEG* is an indicator variable set to 1 if consensus declined across an earnings announcement. Δ*CONSENSUS* in Regression (3) is the difference between *CONSENSUS* in annual earnings forecasts issued by analysts in the 45 days before the interim earnings announcement and *CONSENSUS* in annual earnings forecasts issued by the same analysts in the 30 days after the interim earnings announcement.[9]

---

[7] Even in the unlikely event where individual forecast accuracy ($1/V$) is held constant and the relative degree to which information is private (D/V) increases, the mean forecast still moves closer to realized earnings. Thus, increases in private information always suggest some type of increase in forecast accuracy either in terms of more accurate indvidual forecasts or a more accurate mean forecast or both.

[8] Note: Our empirical measure *PPRECISION* is equivalent to KV theoretical construct only under certain restrictive theoretical assumptions as described in BKLS. In contrast, consensus reflects the HV theoretical construct under a much less restrictive set of theoretical assumptions.

[9] Barron et al. (2002) show that the decrease in consensus after earnings announcements is roughly the same using (1) a shorter 30-day pre-announcement window and a 10-day post-announcement window and (2) a longer 45-day pre- and 30-day post-announcement windows. This similarity suggests that changes in BKLS consensus measured with the longer windows are primarily attributable to the effects of earnings announcements. Because we require a minimum of five analysts to revise their forecasts, using the wider window expands the sample.

We test the relation between change in consensus and residual volume in two ways for the following reasons. Consistent with the HV theory's focus on decreases in consensus, in Regression (2) we use an indicator variable set to 1 if consensus declined across an earnings announcement period, *CONNEG*. This approach tests whether earnings announcements that reduce consensus stimulate more trading volume than announcements that increase consensus. This dichotomous specification is relatively insensitive to influential observations, measurement error, and misspecifications such as nonlinearity. However, because the indicator variable likely removes some information, we re-estimate Regression (3) with a continuous variable measuring the change in consensus (and as indicated in our descriptive statistics consensus increases for approximately 40 percent of our sample).

Consistent with the argument that declines in consensus likely reflect increases in private information rather than decreases in common information, BKLS consensus is negatively related to BKLS precision of analysts' private information. Specifically, when all analysts have the same quality of private information, BKLS consensus can be written as the ratio of common to total (common plus private) information:

$$\rho = \frac{h}{h + s} \tag{6}$$

where h is the precision of all analysts' common information and s is the precision of each analyst's private information. As is obvious from inspection of Equation (6), an increase in private event-period information, s, is not the only possible explanation for a decrease in consensus, ρ. A decrease in BKLS consensus could also occur because the announcement reduced the amount of common information available, h. However, it is not very plausible that earnings announcements routinely reduce the total *amount* of common information, h. Indeed, Barron et al. (2002) show that, on average, earnings announcements increase the amount of common information. Moreover, consistent with our conjecture that decreases in consensus are more likely attributable to increases in s than to decreases in h, we find that the Spearman correlation between $\Delta PPRECISION$ and $\Delta CONSENSUS$ for our sample is -0.514, which is significant at 0.0001. Also supporting this conjecture, Botosan et al. (2004) find a negative correlation between analysts' consensus and the precision of their private information during a later time period, 1993 to 2001. These findings support our interpretation of decreases in consensus as an indicator of increases in private information around earnings announcements.

Control Variables

LARETURN

We abstract from trading attributable to value-relevant information in the earnings announcement by including the natural log of the absolute value of firm returns during the three-day event window. This controls for trading volume associated with price changes resulting from the earnings announcement, one of the types of information-based trading in KV.

LJUMBLING

We include the measure of belief reordering, or jumbling, used by Barron (1995) and Bamber et al. (1997) to control for trading attributable to differences in the precision of investors' private pre-disclosure information, another type of information-based trading in KV. Both studies define "belief jumbling" as differential belief revisions that change an individual's expectation relative to the distribution of expectations held by others—in other words, the reordering of beliefs (or forecasts) across investors (or analysts). We measure belief jumbling as the complement of the correlation between the relative positions of individuals' pre- and post-announcement forecasts. Bamber et al. (1997) contend that the volume associated with belief jumbling around earnings announcements likely reflects the extent to which announcements trigger disagreement about a security's value attributable to differences in the precision of individual investors' private pre-disclosure information.

A limitation of this belief jumbling measure is that it captures more than just the differential precision of analysts' private pre-disclosure information. Specifically, as noted in Kandel and Pearson (1995), belief jumbling also captures belief revisions that violate the usual economic assumption that rational agents with identical information will have similar expectations. Fortunately, we seek to control for this form of disagreement as well because it likely leads to some trading unassociated with the development or use of private event-period information.

Nonetheless, some jumbling may be caused by reductions in consensus that reflect the development of private event-period information. To the extent that controlling for jumbling abstracts from part of the effect of private event-period information, this weakens our tests and biases against finding the hypothesized results. However, we expect a minimal correlation, if any, between jumbling and our proxy for event-period private information, *CONSENSUS*, because consensus is the correlation in forecast errors at *a point in time*, while belief jumbling is the correlation *over time* between pre- and post-forecasts. Any given correlation in forecast errors at a point in time is consistent with many different relative positions of individual forecasts within the entire distribution of forecasts. As a result, significant changes in the relative positions of individual forecasts (or belief jumbling) may occur without changing the overall correlation in forecast errors (i.e., forecasts may jumble without changing forecast dispersion [*D*] or the variance of forecasts errors [*V*]). Likewise, significant changes in the degree to which forecast errors are correlated at a point in time (consensus) need not be associated with changes over time in the relative positions of individual forecasts within the distribution of forecasts (jumbling). That is, forecasts may converge and exhibit a more common error without becoming reordered (or jumbled). Indeed, as shown later in Table 2, in our sample belief jumbling is not significantly correlated with either Δ*CONSENSUS*, *CONNEG*, or Δ*PPRECISION*. Thus, controlling for jumbling is unlikely to abstract from the effect of interest—trading volume induced by the development of private event-period information.

In summary, we test H1 and H2 by testing for an association between: (1) excess trading volume and (2) either (a) changes in the BKLS proxy for the precision (amount) of analysts' private information or (b) changes in the BKLS proxy consensus. Our tests control for trading associated with price changes and the Bamber et al. (1997) proxy for differences in the precision of analysts' pre-disclosure private information. A positive association between trading volume and our proxy for the precision of analysts' private information would support the KV prediction that some investors develop or acquire private event-period information that affects their trading. A negative relation between trading volume and changes in consensus would support the HV model of trading volume. Finding both relations would strengthen the inference that investors develop or acquire private event-period information around earnings announcements, which is sufficiently important to influence their actions (i.e., their trading).

IV. THE SAMPLE AND EMPIRICAL EVIDENCE

This section describes the sample and reports on the results of our tests. The sample consists of firm-quarter observations for quarters one through three between the first quarter of 1984 and the last quarter of 1996,[10] inclusive, that meet the following criteria:

1. The quarterly (interim) earnings announcement date is available on the quarterly Compustat Primary, Supplementary, or Tertiary file;
2. Five or more individual analysts on I/B/E/S issued annual earnings forecasts within a 45-day pre-announcement window and then revised their forecast within a 30-day post-announcement window; and
3. Trading volume and stock return data are available on the CRSP NYSE, AMEX, and NASDAQ files.

A total of 2,977 firm-quarter observations meet criteria (1) and (2), though 116 are missing volume or returns data. In supplemental analyses we compare our findings to evidence in earlier empirical research that used different proxies for investor disagreement. Because the calculation of these previously used variables requires deflating by the mean earnings-per-share forecast, we also deleted 137 observations with earnings-per-share or mean annual earnings forecasts of less than 0.10 in absolute value (to avoid small denominator problems in these tests). The final sample consists of 2,724 firm-quarter observations.

The second sample selection criterion is the most restrictive, but it ensures that pre- and post-announcement measures of consensus are based on *recent* forecasts by the same analysts both before and after earnings announcements. As Bamber et al. (1997, 1999) point out, this requirement increases the precision of forecast-based proxies while helping to ensure that test results are uninfluenced by stale forecasts. This criterion also helps to ensure that we are primarily capturing changes in analyst consensus attributable to the quarterly earnings announcement. One limitation is that this criterion leads to a sample of relatively large firms with large quarterly earnings surprises.

Descriptive Statistics

Table 1 presents descriptive statistics for the sample. As evidenced by the large price reaction, the sample firms have relatively large quarterly earnings surprises. The mean (median) absolute return (*ARETURN*) for the three days surrounding the quarterly earnings announcement is 0.061 (0.046). In addition, the sample firms are quite large, the mean (median) market value of equity is $7,169 ($2,950) million, which is in the ninth decile of market value for all Compustat firms. Again, this result probably reflects the requirement of five forecast revisions.

For descriptive purposes, we report summary statistics on both the percentage of outstanding shares traded over the three days surrounding the earnings announcement (*VOL*) and our measure of excess trading volume (*EXVOL*) that controls for non-information based trading. The mean (median) of *VOL* in this event window is 2.04 percent (1.15 percent), a result comparable to the 1.54 percent (1.02 percent) Bamber et al. (1997) reported for a sample subject to similar restrictions from an earlier (but overlapping) time period. The dependent variable, *EXVOL*, controls for non-information-based trading by subtracting the log of firm-specific median volume (defined as the median amount of volume for contiguous three-day periods over the 249 trading days prior to the earnings announcement event window) from the log of the event-period volume. *EXVOL* has a mean (median) of 0.384 (0.345). Unlike prior measures of excess volume that are positively skewed, *EXVOL* is

---

[10] This is a representative time period that avoids the technology stock price run-up and crash.

TABLE 1
Descriptive Statistics for a Sample of 2,724 Interim Earnings Announcements over the Period 1984 to 1996[a]

Variable	Mean	25th Percentile	Median	75th Percentile
VOL (% traded)	2.04%	0.66%	1.15%	2.14%
EXVOL	0.384	0.005	0.345	0.737
ΔPPRECISION	278,647	0.011	16.5	647.6
CONNEG	0.604	0	1	1
ΔCONSENSUS	−0.082	−0.234	−0.026	0.046
Ranked *ΔPPRECISION*	463.53	227.5	454.5	681.5
Control Variables				
ARETURN	0.061	0.029	0.046	0.074
LARETURN	−3.081	−3.526	−3.070	−2.600
LJUMBLING	−1.018	−1.563	−0.845	−0.257
SIZE ($ millions)	$7,169	$1,345	$2,950	$7,253

[a] Each interim announcement has at least five analysts' forecasts of annual earnings updated within 45 days before the quarterly (interim) announcement, with the same five analysts updating within 30 days after the announcement.

Variable Definitions:

VOL = the percent of outstanding shares traded over the three-day event window (days −1 to +1 relative to the quarterly earnings announcement date);

EXVOL = event-period induced excess trading volume, defined as the natural log of the cumulative trading volume over the three-day event window (days −1 to +1 around the interim earnings announcement) minus the natural log of the firm-specific median trading volume, where median volume is defined as the median amount of volume for contiguous three-day periods over the 249 trading days prior to the earnings announcement event window;

ΔPPRECISION = a continuous variable measuring changes in the BKLS proxy for the precision (or amount) of analysts private information;

CONNEG = an indicator variable set to 1 if the BKLS proxy for analysts' consensus declined around the earnings announcement, and 0 otherwise;

ΔCONSENSUS = a continuous variable measuring changes in the BKLS proxy for analysts' consensus around the earnings announcement;

ARETURN = the absolute value of the sum of a firm's daily returns in the three-day event window;

LARETURN = the natural log of the absolute value of the sum of firm's daily returns in the three-day event window;

LJUMBLING = the natural log of 1 minus the Pearson correlation between the annual earnings forecasts issued by individual analysts in the 45 days before the interim earnings announcement and the annual earnings forecasts issued by the same analysts within the 30 days after the interim earnings announcement; and

SIZE = market value of equity, measured six days prior to the three-day event window.

approximately normally distributed, because it is the difference between two logged measures of trading volume.

We use Equation (4) to calculate the proxy for the precision of analysts' private information. As discussed previously, *PPRECISION* is not naturally scaled and has extreme values. Accordingly, to test H1 we estimate Equation (1) using rank regression. Consistent with scale problems attributable to differences in the magnitude of error across firms or the fact that *PPRECISION* is unbounded, the mean change in *PPRECISION* is 278,647 and the median is 0.165. The ranked values of change in *PPRECISION* are much better behaved, with a mean (median) of 463.5 (454.5). A rank regression is used in Regression (1) with each variable ranked within each quarter, due to slightly different numbers of observations in each quarter (868 in quarter 1, 775 in quarter 2, and 1081 in quarter 3) the mean and median ranked values of *PPRECISION* are not equal. A weakness of this proxy is that it may be subject to measurement error caused by scale problems not fully accounted for by rank regression.

We use Equation (5) to calculate the proxy for analysts' consensus. Consensus declines for 60.4 percent of the sample firms, the mean of *CONNEG*. The mean (median) change in analysts' consensus, Δ*CONSENSUS*, is −0.082 (−0.026). Even with this broader sample of earnings announcements, the result is similar to the declines in consensus Barron et al. (2002) report. In contrast to *PPRECISION*, the consensus proxies are naturally scaled (between −1 and 1) because *CONSENSUS* reflects the amount of private information relative to other information.

The control variables are *LARETURN* and *LJUMBLING*. The mean (median) of *LARETURN* is −3.08 (−3.07). The mean (median) of our unlogged jumbling variable is 0.538 (0.429), which is comparable to the 0.477 (0.589) that Bamber et al. (1997) reported.

Pairwise Correlations

Table 2 presents Spearman (above diagonal) and Pearson (below diagonal) correlations among all variables included in our analyses. Because they are similar, we discuss the Pearson correlations. The correlation between excess trading volume and the ranked change in the precision of analysts' private information is 0.030, which is insignificant (p-value of 0.1203). However, quarterly earnings announcements that decrease consensus about upcoming annual earnings (*CONNEG* = 1) spur more trading volume than do announcements that increase consensus (p-value of 0.0001) and change in consensus correlates negatively with excess volume with a correlation coefficient of −0.064 (p-value of 0.0008). Untabulated analyses reveal the median announcement period increase in volume to be 38 percent greater for firms with a decline in consensus than it is for firms with an increase in consensus. These results provide preliminary evidence consistent with the hypothesis that those

TABLE 2

Spearman (above diagonal) and Pearson (below diagonal) Correlations among Excess Trading Volume, Ranked Change in Private Event-Period Information, Declines and Changes in Analysts' Consensus, and Control Variables[a]

(n = 2,724; p-values in parentheses)

Variable	*EXVOL*	**Ranked** Δ*PPRECISION*	*CONNEG*	Δ*CONSENSUS*	*LARETURN*	*LJUMBLING*
EXVOL		0.031 (0.1008)	0.084 (0.0001)	−0.065 (0.0007)	0.445 (0.0001)	0.090 (0.0001)
Ranked Δ*PPRECISION*	0.030 (0.1203)		0.390 (0.0001)	−0.469 (0.0001)	0.039 (0.0400)	0.019 (0.3251)
CONNEG	0.092 (0.0001)	0.379 (0.0001)		−0.847 (0.0001)	0.080 (0.0001)	0.013 (0.5078)
Δ*CONSENSUS*	−0.064 (0.0008)	−0.504 (0.0001)	−0.664 (0.0001)		−0.071 (0.0002)	−0.018 (0.3351)
LARETURN	0.429 (0.0001)	0.037 (0.0516)	0.079 (0.0001)	−0.058 (0.0025)		0.100 (0.0001)
LJUMBLING	0.090 (0.0001)	0.028 (0.1508)	0.018 (0.3506)	−0.021 (0.2788)	0.081 (0.0001)	

[a] Table 1 describes the sample and variables.

quarterly earnings announcements that reduce the consensus about upcoming annual earnings lead to more announcement-induced trading volume. In addition, consistent with increases in private information reducing consensus, the correlation between *CONNEG* and *ΔPPRECISION*, is significantly positive (p-value of 0.0001). Similarly, the correlation between *ΔCONSENSUS* and *ΔPPRECISION* is significantly negative, −0.469 (p-value of 0.0001).

The control variable for information-induced volume related to price changes is *LARETURN*. Consistent with this variable acting as a control, the correlation between excess volume and *LARETURN* is 0.429 (p-value of 0.0001). Similarly, the correlation between *LJUMBLING* and *EXVOL* is 0.090 (p-value of 0.0001).

We assume that the jumbling proxy primarily reflects the extent to which earnings announcements trigger new-found disagreement about the value of a security attributable to differences in the precision of individual investors' private pre-disclosure information. Although, jumbling may also reflect a reduction in consensus attributable to private event-period information, insignificant correlations between jumbling and both measures of changes in consensus (p-values of 0.2788 and 0.3506, respectively) suggest that in our sample jumbling does not reflect changes in consensus. In addition, in models such as KV, belief jumbling increases with both differences in precision and the magnitude of price changes. Consistent with this, in our sample the correlation between *LJUMBLING* and *LARETURN* is significantly positive (p-values of 0.001). These results are consistent with our interpretation of the jumbling variable as a control for differential precision of private pre-disclosure information.

Results of Hypothesis Tests

To address the dependency that results from including in the sample multiple observations from the same firm, we calculate the t-statistics reported in Tables 3 and 4 using the Froot (1989) procedure. Table 3 reports the results of testing H1 using rank regression analysis to estimate Equation (1). Our proxy for KV private event-period information is the

TABLE 3
Rank Regression of Excess Trading Volume around Quarterly Earnings Announcements on Changes in the Precision of Analysts' Private Information and Control Variables
(n = 2,724)

Variable[a]	Expected Sign	Coefficient (p-value)[b]
Intercept		193.7 (0.0001)
ΔPPRECISION	+	0.047 (0.0054)
LARETURN	+	0.069 (0.0001)
LJUMBLING	+	0.466 (0.0001)
Adjusted R^2		23.7%

[a] The notes to Table 1 describe the sample and variables.
[b] p-values are one-sided where tests predictions are made and are calculated using the Froot (1989) method.

precision (amount) of analysts' private information. Consistent with H1, the coefficient on $\Delta PPRECISION$ is significantly positive, 0.047 (p-value of 0.0054). This finding suggests that earnings announcements that increase analysts' private information also induce investors to trade in a manner consistent with an increase in their private information.[11] As expected, the coefficients on the control variables for price changes (*LARETURN*) and differential precision of pre-disclosure information (*LJUMBLING*) are both positive and significant, consistent with prior research (Bamber et al. 1997).

Table 4 reports the results of testing H2 by estimating Equations (2) and (3). The positive *CONNEG* coefficient (p-value of 0.0003) is consistent with the hypothesis that earnings announcements that decrease consensus are associated with more excess trading volume than earnings announcements that increase consensus.[12] In addition, in the estimation of Equation (3), the coefficient on the continuous change-in-consensus variable $\Delta CONSENSUS$ is negative (p-value of 0.0137). These results suggest that the more a quarterly earnings announcement decreases consensus about the upcoming annual earnings, the greater the residual trading volume around the earnings announcement. Again, the signs and significance levels of the coefficients on the control variables are as expected. Because decreases in consensus are consistent with some investors having private event-period information, these results suggest that this new private information causes investors to value securities differently and to trade.

TABLE 4

Regression of Excess Trading Volume around Quarterly Earnings Announcements on Changes and Declines in Analysts' Consensus, and Control Variables

(n = 2,724)

Variable[a]	**Expected Sign**	**Coefficient (p-value)**[b]	
Intercept		1.330 (0.0001)	1.372 (0.0001)
CONNEG	+	0.069 (0.0003)	
ΔCON	−		−0.067 (0.0137)
Controls:			
LARETURN	+	0.310 (0.0001)	0.312 (0.0001)
LJUMBLING	+	0.031 (0.0003)	0.031 (0.0003)
Adjusted R^2		18.76%	18.94

[a] The notes to Table 1 describe the sample and variables.

[b] p-values are one-sided where tests predictions are made and are calculated using the Froot (1989) method.

---

[11] Supplementary analysis revealed that BKLS measures of changes in the precision of *public* information are *not* significantly associated with residual trading volume following earnings announcements. Thus, our findings are not simply due to the fact that analysts' and investors' total amount of information increases around earnings announcements.

[12] We examined the effects of extreme values on our results by winsorizing our dependent variable and by deleting influential observations (two passes). We also estimated models (2) and (3) using rank regression similar to model (1). All inferences remain unchanged.

V. SUPPLEMENTAL ANALYSES

Prior empirical research has found that other variables are associated with excess trading volume around earnings announcements. Consequently, we investigate whether changes in the proxies for the amount of private information and consensus continue to be important indicators of earnings-announcement-induced trading volume after controlling for these other variables.

First, as a proxy for market-wide trading volume, we control for the natural log of the percentage of shares traded over the three-day event window for all firms on the CRSP NYSE, AMEX, and NASDAQ files. Second, we include changes in the dispersion of analysts' forecasts, measured as the change in the standard deviation of annual earnings forecasts issued 45 days before and 30 days after the interim earnings announcement, scaled by the absolute value of the mean annual earnings forecast.[13] Untabulated analyses reveal that the coefficient on $\Delta PPRECISION$ remains significantly positive (p-value of 0.0307). Similarly, the coefficient on *CONNEG* remains significantly positive and the coefficient on $\Delta CONSENSUS$ significantly negative (p-values of 0.001 and 0.052, respectively). Thus, including these additional control variables does not affect our inferences. Consistent with prior research, the coefficients on market-wide volume and change in dispersion are positive with p-values less than 0.051 (Bamber et al. 1997; Ziebart 1990).[14] Even though the same forecasts are used to construct $\Delta CONSENSUS$, $\Delta PPRECISION$, belief jumbling, and change in forecast dispersion, both $\Delta CONSENSUS$ and $\Delta PPRECISION$ retain incremental explanatory power, consistent with the HV and KV theoretical models.

We also added to the regression models two variables that Bamber (1987) showed to be related to trading volume around earnings announcements. First, as a direct measure of earnings surprise, we control for the absolute value of the difference between the mean analysts' forecast and actual earnings per share (before extraordinary items), deflated by the absolute value of the mean forecast. Second, we also control for firm size, measured as the market value of equity at the end of day −6 relative to the earnings announcement. Following prior research, we log-transformed both of these variables. When these variables are added to the regression model, the coefficient on $\Delta PPRECISION$ remains significantly positive (p-value of 0.0029). The coefficients on earnings surprise and firm size also remain significantly positive with p-values of 0.0021 and 0.0003, respectively. Similarly, the finding that decreases in consensus stimulate trading remained unchanged with p-values of less than 0.008 for both coefficients. Both earnings surprise and firm size are, however, insignificant when added to the regression models with $\Delta CONSENSUS$.[15]

To test whether our findings are robust across differing levels of unexpected earnings, we also repeated our analysis on subsamples based on quartiles of the magnitude of unexpected earnings.[16] The results for $\Delta PPRECISION$ are mixed—the coefficient is significantly positive only in quartiles two and three. The coefficient on the dichotomous variable,

---

[13] BKLS decompose dispersion into both consensus and overall uncertainty. Including the change in their uncertainty measure in our analyses has no effect on our inferences.

[14] Our supplementary findings concerning changes in forecast dispersion are consistent with Barron and Karpoff's (2004) theoretical prediction that, after controlling for consensus effects, transaction costs prevent more trades when beliefs become less dispersed than when beliefs become more dispersed.

[15] In the model with $\Delta PPRECISION$, the coefficients on earnings surprise and size are significantly positive, consistent with prior research. However, in the models with consensus the coefficients on earnings surprise and size are insignificant. Further analysis shows the earnings surprise and firm size effects are subsumed when we include absolute returns in the model.

[16] The results are insensitive to the inclusion of indicator variables that control for whether the earnings announcement pertains to the first, second, or third fiscal quarter.

CONNEG, is significantly positive across all four quartiles but the coefficient on the continuous measure, $\Delta CONSENSUS$, although consistently of the hypothesized sign, is significantly negative only in quartiles two and four. Measurement error or some unidentified nonlinearity may vary across the quartiles. Alternatively, there may be insufficient intra-quartile variation in the magnitude of earnings surprise in our sample because our firms all evidence relatively large earnings surprises.

Conference calls to selected analysts could result in private event-period information of a type that is not the focus of this study (that is private information that is not the result of "interactive" knowledge obtained prior to the announcement that can only be capitalized upon after seeing the announcement). Tasker (1997, 8) reports that conference calls to analysts around earnings announcements were very uncommon before the mid-1990s. Consequently, we repeated our tests on the subsample of 1,226 earnings announcements occurring prior to 1991. The results are qualitatively similar to those for the larger sample, which provides evidence that the primary findings are not the result of conference calls.[17]

In summary, proxies for the theoretically motivated variables—precision of analysts' private information and consensus—remain significant even after controlling for other, less theoretical, determinants of trading volume, such as firm size.

VI. SUMMARY AND CONCLUSIONS

Kim and Verrecchia's (1997) theoretical model predicts that private information inferred at the time of an information release (here, an earnings announcement) is associated with greater trading volume. We provide empirical evidence that announcements that lead to increases in analysts' private information (as measured by Barron et al.'s [1998] empirical proxies) also lead to increases in trading volume suggesting that some investors, as well as analysts, develop or acquire private event-period information. In addition, consistent with Holthausen and Verrecchia's (1990) theoretical model, decreases in consensus lead to increases in trading volume. Because consensus declines when private information increases, this finding is consistent with theory and provides reinforcing evidence that earnings announcements lead some investors to trade based on private event-period information.

Our analyses also reveal that our theoretically derived measures of analysts' private information and consensus remain significant after controlling for other determinants of trading volume documented in prior research (e.g., change in dispersion of analysts' forecasts, belief jumbling, earnings surprise, and firm size). Overall, our results support the theoretical proposition that part of the trading around earnings announcements arises because investors use new private information developed at the time of the announcement in addition to prior and contemporaneous value-relevant information.

Although explaining more of the trading volume around earnings announcements is important simply because it is the most detectable indicator that financial accounting influences individual investors' decision making and actions (Cready and Hurtt 2002), one can best appreciate the importance of our findings by viewing them together with the findings of other accounting research concerning firms' costs of capital. Botosan et al. (2004) find analysts' private information is positively associated with a firm's cost of capital, and Barron et al. (2002) find that earnings announcements trigger the release of a large percentage of analysts' private information. Kim and Verrecchia (1994, 1997) contend that it is trade based on the type of private information reported by Barron et al. (2002) that adversely affects firms' costs of capital. Thus, when considered together with the findings of others, our results are consistent with the argument that private information triggered by earnings announcements is a factor in determining firms' costs of capital.

---

[17] Earnings management to meet or beat analysts' forecasts adds measurement error to the BKLS proxies for the precision of private information and consensus. Fortunately, however, earnings management to meet or beat analysts' forecasts was also less common prior to 1992 (Richardson et al. 2004). This information is important because this type of earnings management poses a potential threat to inferences derived using the BKLS method.

One caveat centers on the fact that the sample selection criteria led to a sample of large firms (the median firm is in the ninth decile of the Compustat population) with relatively large quarterly earnings surprises. The extent to which our inferences generalize to other populations of smaller firms or less surprising earnings announcements remains an open question.

REFERENCES

Atiase, R., and L. Bamber. 1994. Trading volume reactions to annual accounting earnings announcements: The incremental role of pre-disclosure information asymmetry. *Journal of Accounting and Economics* 17 (May): 309–329.

Bamber, L. 1987. Unexpected earnings, firm size, and trading volume around quarterly earnings announcements. *The Accounting Review* (July): 510–532.

———, O. Barron, and T. Stober. 1997. Trading volume around earnings announcements and different aspects of disagreement. *The Accounting Review* (October): 575–597.

———, ———, and ———. 1999. Differential interpretations and trading volume. *Journal of Financial and Quantitative Analysis* 34 (3): 369–386.

Barron, O. 1995. Trading volume and belief revisions that differ among individual analysts. *The Accounting Review* (October): 581–597.

———, O. Kim, S. Lim, and D. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* (October): 421–33.

———, D. Byard, and O. Kim. 2002. Changes in analysts' information around earnings announcements. *The Accounting Review* (October): 821–46.

———, and J. Karpoff. 2004. Information precision, transaction costs, and trading volume. *Journal of Banking and Finance* (June): 1207–1223.

Beaver, W. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6 (Supplement): 67–92.

Botosan, C., M. Plumlee, and Y. Xie. 2004. The role of private information precision in determining cost of equity capital. *Review of Accounting Studies* (June-September): 233–259.

Cready, W., and D. Hurtt. 2002. Assessing market response using return and volume metrics. *The Accounting Review* (October): 891–909.

Dontoh, A., and J. Ronen. 1993. Information content of accounting announcements. *The Accounting Review* (October): 857–869.

Froot, K. 1989. Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis* (September): 333–355.

Holthausen, R., and R. Verrecchia. 1990. The effect of informedness and consensus on price and volume behavior. *The Accounting Review* (January): 191–208.

Kandel, E., and N. Pearson. 1995, Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103 (4): 831–872.

Kim, O., and R. E. Verrecchia. 1991. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29 (Autumn): 302–321.

———, and ———. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* (January): 41–67.

———, and ———. 1997. Pre-announcement and event-period private information. *Journal of Accounting and Economics* (December): 395–419.

Kross, W., G. Ha, and F. Heflin. 1994. A test of risk clientele effects via an examination of trading volume response to earnings announcements. *Journal of Accounting and Economics* (July): 67–87.

Richardson, S., S. Teoh, and P. Wysocki. 2004. The walkdown to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Ross, S. 1989. Discussion: Intertemporal asset pricing. In *Theory of Valuation*, edited by S. Bhattacharya, and G. Constantinides, 85–96. Totowa, NJ: Rowman & Littlefield.

Tasker, S. 1997. Voluntary disclosure as a response to low accounting quality: Evidence from quarterly conference calls. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Ziebart, D. 1990. The association between consensus of beliefs and trading activity surrounding earnings announcements. *The Accounting Review* (April): 477–488.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 423–440

A Temporal Analysis of Quarterly Earnings Thresholds: Propensities and Valuation Consequences

Lawrence D. Brown
Georgia State University

Marcus L. Caylor
Georgia State University

ABSTRACT: Applying a Burgstahler and Dichev (1997)/Degeorge et al. (1999) type methodology to quarterly data for the 1985–2002 time period, we show that, since the mid-1990s, but not before then, managers seek to avoid negative quarterly earnings surprises more than to avoid either quarterly losses or quarterly earnings decreases. Our findings suggest that the quarterly earnings threshold hierarchy proposed by Degeorge et al. (1999) does not apply to recent years, and that managers' claim that avoiding quarterly earnings decreases is the threshold they most seek to achieve (Graham et al. 2004) is inconsistent with their actions. We provide an intuitively appealing economic rationale for why the shift in threshold hierarchy occurred; since the mid-1990s, but not before then, investors unambiguously rewarded (penalized) firms for reporting quarterly earnings meeting (missing) analysts' estimates more than they did for meeting (missing) the other two thresholds. We provide several explanations for why investors unambiguously reward firms for reporting quarterly earnings that meet or beat analysts' estimates more than for meeting the other two thresholds late (but not early) in our sample period: increased media coverage given to analyst forecasts, more analyst following, more firms covered by analysts, and temporal increases in both the accuracy and precision of analyst forecasts.

Keywords: *thresholds; propensities; valuation consequences; temporal changes; analyst estimates.*

Data Availability: *All data are available from public databases identified in the paper.*

We appreciate the comments and suggestions of Bill Baber, Sudipta Basu, Dave Burgstahler, Jennifer Francis, Leslie Hodder, Debra Jeter, Chris Jones, Sok-Hyon Kang, Krishna Kumar, Tom Linsmeier, Liyu Luo, Siva Nathan, Per Olsson, Julie Petherbridge, Sundaresh Ramnath, Mohan Venkatachalam, participants of seminars at The George Washington University, Georgia State University, Michigan State University, State University of New York at Buffalo, Vanderbilt University, 2003 American Accounting Association Annual Meeting, 2003 Conference on Financial Economics and Accounting, 2003 Southeast Summer Accounting Research Colloquium, 2004 Financial Accounting and Reporting Section Midyear Meeting, and two anonymous reviewers. We are grateful to Thomson Financial I/B/E/S for providing data on analysts' earnings forecasts.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted September 2003
Accepted August 2004

I. INTRODUCTION

Degeorge et al. (1999) examine the relative importance of three quarterly earnings thresholds managers seek to achieve, and conclude that managers order their earnings thresholds as follows: (1) avoid quarterly losses; (2) avoid quarterly earnings decreases, and (3) avoid negative quarterly earnings surprises.[1] Dechow et al. (2003) examine these same thresholds using annual data, and show that avoiding negative annual earnings surprises is the most important threshold managers seek to achieve in the last three years of their sample, 1999–2001, but not before then. In contrast to Dechow et al. (2003), Graham et al. (2004) find that managers maintain that avoiding quarterly earnings decreases is the most important quarterly earnings threshold that they seek to achieve.[2]

These collective findings give rise to several important unanswered questions. Does the managerial threshold hierarchy shown by Dechow et al. (2003) for annual earnings thresholds pertain to quarterly earnings thresholds in recent years? If so, then the hierarchy documented by Degeorge et al. (1999) in their influential study is no longer applicable, and what managers say their priorities are regarding achieving quarterly thresholds (the Graham et al. [2004] evidence) differs from what they actually do.[3] If the threshold hierarchy proposed by Degeorge et al. (1999) is no longer applicable, then when and why did the shift occur? Alternatively, if the Degeorge et al. (1999) finding that avoiding quarterly earnings decreases is sought after relatively more *vis-à-vis* avoiding negative quarterly earnings surprises is still valid, then what managers say they do regarding setting their earnings threshold priorities (the Graham et al. [2004] findings) is consistent with what they actually do, but threshold findings based on annual data (the Dechow et al. [2003] findings) do not apply to quarterly data.

Consistent with Degeorge et al. (1999), we show that managers focus *less* on avoiding negative quarterly earnings surprises than on avoiding *either* quarterly losses or earnings decreases early in our sample period (1985–1993). However, in direct contrast to Degeorge et al. (1999), we show that avoiding negative earnings surprises is not the least important quarterly threshold after 1993. Indeed, it is the *most* important threshold in every year, 1996–2002.

The major reasons managers claim they seek to meet or beat quarterly earnings benchmarks is to build credibility with capital markets and to maintain or increase their firms' stock prices (Graham et al. 2004). Consistent with this view, we assume that managers act in their own best interests, and that their large holdings of stock, options, and fear of losing their jobs if their firms' stock prices drop sharply if they fail to achieve thresholds induce them to look to the capital markets when they consider which threshold to focus upon most intently. We analyze incremental valuation consequences of making (missing) thresholds, and document that since the mid-1990s, but not before then, investors unambiguously reward (penalize) firms more for reporting quarterly earnings that meet (miss) analysts' estimates than for meeting (missing) the other two thresholds. Our findings are consistent with the notion that managers took a cue from the capital markets to re-order their quarterly earnings threshold hierarchy.

---

[1] Using annual data, Burgstahler and Dichev (1997) obtain the same hierarchical ordering for the first two thresholds. Burgstahler and Dichev (1997) do not examine analyst forecasts (the third threshold). Degeorge et al. (1999) conclude that avoiding a loss is the most important threshold to achieve based on their conditional analyses (i.e., making a threshold when another threshold is made/missed), but their unconditional results reveal that avoiding an earnings decrease is the most important threshold. More importantly for the purposes of our study, the Degeorge et al. (1999) finding that avoiding a negative earnings surprise is the least important threshold is robust to their use of both conditional and unconditional methodologies.

[2] Their survey of 312 financial executives from public firms reveals the following rank ordering of achieving earnings benchmarks: (1) same quarter last year (84.3 percent agree); (2) analyst consensus estimate of the current quarter (68.8 percent agree); (3) reporting a profit this quarter (65 percent agree); and (4) previous quarter EPS (54.6 percent agree).

[3] According to the Social Sciences Citation Index, Degeorge et al. (1999) was cited 48 times as of June 22, 2004. As a frame of reference, we compare cites to Degeorge et al. (1999) with cites to all 19 articles published in the first issue of 1999 by the two premier finance journals, *Journal of Finance* and *Journal of Financial Economics* (Trieschmann et al. 2000). Only one article (Barber and Lyon 1999) was cited more often than Degeorge et al. (1999).

After documenting a temporal shift in investor reaction to quarterly earnings, we examine why the shift occurred. We motivate our analyses based on studies showing that valuation consequences of avoiding negative earnings surprises increases in the accuracy (Fried and Givoly 1982; Brown et al. 1987; Wiedman 1996) and precision of analyst forecasts (Imhoff and Lobo 1992; Kinney et al. 2002). We show that, during our sample period, analyst quarterly earnings forecasts have become both more accurate and more precise.

We also document temporal increases in analyst following, firms followed by analysts, and media attention paid to meeting or beating analyst earnings estimates, providing additional evidence for why investors increased their focus on analyst earnings forecasts.

We show that our findings are robust to five sensitivity checks: (1) using a ratio metric suggested by Dechow et al. (2003) in lieu of the Burgstahler and Dichev (1997) measure; (2) un-adjusting earnings data for stock splits; (3) requiring firms to be in the sample for the entire time period; (4) using Compustat actuals in lieu of I/B/E/S actuals; and (5) omitting observations where I/B/E/S earnings differ from GAAP earnings.

Our findings are important for several reasons. First, we provide guidance for which quarterly earnings threshold, academics, practitioners, investors, and regulators should focus on most if they wish to examine earnings management to achieve quarterly earnings thresholds in recent years.[4] Second, we illustrate that the quarterly earnings threshold managers claim to be most important (Graham et al. 2004) differs from what they seek to achieve most, suggesting that what managers say they do differs from what they actually do.[5] Third, we show that our findings based on quarterly data confirm those based on annual data (Dechow et al. 2003) when both studies use similar methodologies, increasing the validity of both studies. Fourth, we provide an intuitively appealing economic rationale for why threshold hierarchies have changed (i.e., managers took a cue from the capital markets, which have made it clear to them that in recent years avoiding negative surprises is more value-relevant than avoiding either losses or earnings decreases), increasing both our understanding of why managers altered their threshold priorities and our confidence in the threshold results. Fifth, we provide several reasons for why investors paid more attention to earnings surprises in recent years, increasing our understanding of why the valuation implications of making/missing thresholds has shifted in favor of avoiding negative earnings surprises (away from avoiding losses and earnings decreases) in recent years.

II. RESEARCH QUESTIONS

Degeorge et al. (1999) conclude that avoiding a negative earnings surprise is the least important of the three *quarterly* earnings thresholds managers seek to achieve. In contrast, Dechow et al. (2003) show that avoiding negative earnings surprises is the most important of the three *annual* thresholds in recent years. Consistent with Degeorge et al. (1999), but in contrast to Dechow et al. (2003), Graham et al. (2004) show that managers maintain that avoiding quarterly earnings decreases is more important to them than avoiding negative *quarterly* earnings surprises. These conflicting findings make it unclear whether use of a Degeorge et al. (1999) type of methodology (used by Dechow et al. [2003] but not Graham et al.[2004]) would show that, in contrast to Degeorge et al., avoiding a negative earnings surprise is the most important *quarterly* earnings threshold in recent years.[6] Our first research question is:

---

[4] Consistent with our findings, Libby and Kinney (2000) and Ng and Tan (2003) conduct experimental studies of auditors, and focus on the "avoid negative earnings surprise" threshold, ignoring both the "avoid loss" and "avoid earnings decrease" thresholds.

[5] Graham et al. (2004, 9) expected to find analyst quarterly earnings forecasts to be the single most important threshold: "Before administering the survey, we expected the analyst consensus estimate to be the most important earnings benchmark. However, the responses indicate that same quarter last year's EPS seems to dominate the analyst consensus estimate."

> **RQ$_1$:** Is avoiding negative quarterly earnings surprises the threshold that managers seek most to achieve in recent years *vis-à-vis* avoiding quarterly losses or quarterly earnings decreases?

After showing that managers have made avoidance of negative quarterly earnings surprises their most sought-after threshold in recent years (but not early in our sample), we appeal to economic intuition to examine why this phenomenon has occurred. We assume that managers' wealth (e.g., stock holdings, option holdings, and job security) is closely tied to their firms' stock prices so that managers alter their threshold hierarchy if the relative incremental valuation consequences of achieving thresholds change. If avoiding negative earnings surprises is the threshold with the greatest incremental valuation consequences in recent years (but not earlier in our sample period), then managers will alter their threshold prioritization in recent years in favor of avoiding negative earnings surprises. Our second research question is:

> **RQ$_2$:** Have the valuation consequences shifted to reward (penalize) managers more for avoiding (reporting) negative quarterly earnings surprises than for avoiding (reporting) quarterly losses or quarterly earnings decreases in recent years?

After showing that avoiding negative earnings surprises is the threshold with the greatest incremental valuation consequences in recent years, we examine why investors reward (penalize) firms more for avoiding (reporting) negative quarterly earnings surprises *vis-à-vis* avoiding (reporting) losses or earnings decreases in recent years. Investors place more weight on more accurate earnings forecasts (Fried and Givoly 1982; Brown et al. 1987; Wiedman 1996). If analyst forecasts have increased in their relative accuracy *vis-à-vis* those of a seasonal random walk model of quarterly earnings, then investors should increase their focus on avoiding negative quarterly earnings surprises relative to that of avoiding quarterly earnings decreases.[7] Our third research question is:

---

[6] Dechow et al. (2003) use a Degeorge et al. (1999) type methodology, albeit with annual data. The Degeorge et al. (1999) type methodology is similar to the one originated by Burgstahler and Dichev (1997). Both Degeorge et al. (1999) and Graham et al. (2004) use quarterly data but, unlike Degeorge et al., Graham et al. adopt a survey approach. Quarterly earnings thresholds may differ from annual earnings thresholds for at least two reasons: (1) quarterly earnings are un-audited while annual earnings are audited; and (2) quarterly earnings thresholds present managers with short-term targets creating a higher-pressure environment *vis-à-vis* annual earnings thresholds (Eccles et al. 2001). Managers may be better able to manage earnings upward shortly before quarterly *vis-à-vis* annual earnings announcements because quarterly reports are unaudited, making it easier for managers to avoid quarterly losses and quarterly earnings decreases. In contrast, managers may be less able to manage annual earnings shortly before they are announced so they focus more on managing analyst expectations, allowing them to avoid negative earnings surprises but not necessarily to avoid losses or earnings decreases.

[7] A seasonal random walk model is defined as earnings changes, the threshold Degeorge et al. (1999) and Graham et al. (2004) find to be more sought after than analyst-based earnings surprises.

RQ_3: Has the relative accuracy of analyst quarterly earnings forecasts *vis-à-vis* a seasonal random walk model increased in our sample period?

Investor reaction to earnings thresholds may have changed in favor of earnings surprises due to increased precision in analyst forecasts. Imhoff and Lobo (1992) and Kinney et al. (2002) show that investors place more weight on forecasts that are more precise, so we examine if the precision of analysts' earnings forecasts has increased over time. If analysts' earnings forecasts have become more precise over time, then investors should focus more on them later than earlier in our time period. Our final research question is:

RQ_4: Has the precision of analysts' quarterly earnings forecasts increased during our sample period?

III. DATA AND METHODOLOGY

Sample Selection

We obtain actual quarterly earnings, analyst forecasts of quarterly earnings, stock prices, and earnings announcement dates from Thomson Financial I/B/E/S for 19 years, 1984–2002. We require earnings lagged by four quarters to measure our second threshold, earnings changes, so our study period is 18 years, 1985–2002. We calculate three-day market-adjusted cumulative abnormal returns around quarterly earnings announcements and obtain SIC codes using CRSP data.

We use the single most recent forecast made by the timeliest analyst(s) prior to the earnings announcement date as our analyst forecast.[8] We define earnings changes as reported earnings in quarter t minus those in quarter $t-4$, and earnings surprises as reported earnings in quarter t minus the analyst's forecast of earnings in quarter t. We deflate earnings levels, earnings changes and earnings surprises by beginning of quarter t stock price.[9] All firm-quarter observations require reported earnings in both quarter t and quarter $t-4$, an analyst forecast of quarter t earnings, stock price at the beginning of quarter t, and returns data for the three days, -1 to $+1$, where day 0 is the earnings announcement day.

Consistent with Burgstahler and Dichev (1997), we delete firms with SIC codes 4400–5000 (regulated industries) and 6000–6500 (financial institutions). To calculate analyst forecast dispersion, we take the standard deviation of analyst forecasts prior to the earnings announcement date, deflated by stock price (Kinney et al. 2002). For this analysis, we require firms to be followed by at least three analysts. To remove the effects of outliers, we delete three-day market-adjusted cumulative abnormal returns in the two extreme percentiles.[10] Our sample for the main analysis consists of 111,485 firm-quarter observations.

---

[8] When there is more than one analyst forecast on the most recent day, we take the mean of the distribution of analyst forecasts. We use the most recent forecast in lieu of the consensus forecast for several reasons. First, it is common in the literature to use the most recent forecast as a measure of expectations (Brown 2001; Bartov et al. 2002). Second, managers tend to pre-announce all the bad news but about half the good news (Soffer et al. 2000), so using the most recent forecast mitigates the confounding effects of earnings pre-announcements. Third, the most recent forecast is more closely related than is the mean consensus forecast to the stock price reaction to earnings announcements (Brown and Kim 1991) so it is a superior measure of the market's earnings expectation.

[9] Burgstahler and Dichev (1997) divided net income by beginning of year market value of common equity from Compustat, where income approximately equals earnings per share times the number of common shares, and market value of common equity approximately equals price per share times the number of common shares. We obtain qualitatively similar results when we deflate earnings levels and earnings surprises by beginning of quarter t stock price and earnings changes by beginning of quarter $t-4$ stock price.

[10] We obtain qualitatively similar results when we do not delete the two extreme percentiles.

Methodology

Our primary analysis uses a statistical test developed by Burgstahler and Dichev (1997) to examine managerial propensity to achieve thresholds.[11] For each interval, we calculate the actual number of observations minus the expected number of observations, divided by the estimated standard deviation of this difference.[12] We compute a standardized difference of "the missing data" just to the left of zero for each threshold. For example, for the "avoid loss" threshold, we compare very small negative earnings levels (i.e., losses) with very small positive earnings levels (i.e., profits) to determine how many very small losses have been "shifted rightward" to become very small profits.[13] Consistent with Burgstahler and Dichev (1997), we consider negative standardized differences with absolute magnitude of 2.33 or more to be threshold achievement (ρ-value = 0.01 in a normalized distribution, one-tailed test), and the threshold with the largest negative standardized difference to be the most pervasive form of threshold achievement.[14] We estimate an OLS regression of standardized differences on year to assess temporal changes in the managerial propensity to achieve thresholds.[15]

To examine valuation consequences of achieving thresholds, we regress market-adjusted cumulative abnormal returns (CAR) in the three days around quarterly earnings announcements on the eight mutually exclusive and collectively exhaustive categories of making/missing thresholds dummy variables plus interaction variables for each of these dummy variables with unexpected earnings:[16]

$$\begin{aligned} CAR = {} & \Sigma Category\ dummy + \beta_1{*}(T1^- \, T2^- \, T3^-){*}UE + \beta_2{*}(T1^+ \, T2^- \, T3^-){*}UE \\ & + \beta_3{*}(T1^- \, T2^+ \, T3^-){*}UE + \beta_4{*}(T1^- \, T2^- \, T3^+){*}UE \\ & + \beta_5{*}(T1^+ \, T2^+ \, T3^-){*}UE + \beta_6{*}(T1^+ \, T2^- \, T3^+){*}UE \\ & + \beta_7{*}(T1^- \, T2^+ \, T3^+){*}UE + \beta_8{*}(T1^+ \, T2^+ \, T3^+){*}UE + \text{u}. \end{aligned} \tag{1}$$

---

[11] To validate our calculation of the Burgstahler and Dichev metric, we replicated Burgstahler and Dichev (1997) using annual data and their time period and obtained nearly identical results.

[12] The expected number of observations in an interval is assumed to be the average of the two immediately adjacent intervals. The variance is approximately $Np_i(1 - p_i) + (1/4)\,N\,(p_{i-1} + p_{i+1})\,(1 - p_{i-1} - p_{i+1})$ where N is the total number of observations and p_i is the probability an observation falls into interval i. Burgstahler and Dichev (1997) provide details regarding the first two thresholds; Burgstahler and Eames (2003) provide details for the third threshold. We report results using an interval width of .0025, but our general inferences are the same with interval widths of .005 or .01.

[13] We include zeros as achieving thresholds in our primary analyses. However, zeros are biased toward on-target earnings announcements more than those that actually occur (Baber and Kang 2002) so we omit zeros in the "Split-Adjusted Data" section below when we conduct sensitivity analyses.

[14] The Burgstahler and Dichev test statistic represents the deficiency of expected observations in the bin just to the left of zero, indicating that observations have been shifted to the bin just to the right of (and including) zero. Although we do not report Burgstahler and Dichev test statistics just to the right of zero, untabulated results confirm that positive Burgstahler and Dichev test statistics for the bin just to the right of (and including) zero have increased significantly over time for the avoid negative earnings surprise threshold.

[15] The Burgstahler and Dichev test statistic increases in sample size (see Burgstahler and Eames [2003, 292, fn 14]) so, after calculating the distribution across bins using the exact sample, we set all sample sizes equal to 1,000 for all years and thresholds and use the actual proportion of observations in a bin to determine the number of observations given the sample size was 1,000. This grosses down the Burgstahler and Dichev test statistic, providing a conservative bias since our threshold/year sample sizes always exceed 1,000. More specifically, the avoid a loss threshold has mean (median) annual sample sizes of 6,130 (6,283) with a range of 1,952 to 9,461. The avoid an earnings decrease threshold has mean (median) annual sample sizes of 6,114 (6,266) with a range of 1,925 to 9,458. The avoid negative surprise threshold has mean (median) annual sample sizes of 6,161 (6,298) with a range of 1,962 to 9,483. We get similar results with sample sizes of 2,000 or 3,000. More specifically, the t-values of temporal regressions, R^2s, and point where thresholds cross over are identical for sample sizes of 1,000, 2,000, or 3,000.

[16] For simplicity, we omit the i and t subscripts for the dependent and independent variables.

T1, *T2*, and *T3* designate the avoid losses, earnings decreases, and negative surprise thresholds, respectively; the superscripts, $-/+$, indicate if a threshold has been missed or made; ΣCategory dummy are categories of making/missing thresholds dummy variables; and *UE* is unexpected earnings, defined as actual earnings minus the consensus analyst forecast (deflated by beginning of quarter *t* stock price). Each category is represented by a dummy variable (1 if the firm-quarter *t* observation falls into that category, and 0 otherwise). For example, if a firm makes zero (all three) thresholds, the $T1^- \ T2^- \ T3^-$ ($T1^+ \ T2^+ \ T3^+$) dummy variable equals 1. The eight categorical dummy variables collectively represent the intercept of our model (i.e., the eight dummy vectors sum to the unit vector).[17] Each dummy variable is interacted with the magnitude of unexpected earnings deflated by stock price (an earnings response coefficient).

To provide evidence on the incremental valuation consequences of achieving thresholds, we estimate an OLS regression of Equation (1) each year and obtain dummy and interaction coefficients. We calculate the average valuation consequence of making/missing a threshold by summing the value of the dummy variable for that threshold/year and the product of the appropriate interaction term coefficient and the mean level of deflated unexpected earnings for that threshold combination category for that year. We determine incremental valuation consequences by taking differences of these valuation consequences. For example, to obtain the positive incremental valuation consequences for making the avoid loss threshold when no threshold has been made, we subtract ((dummy variable for $T1^- \ T2^- \ T3^-$) + β_1*Mean UE*) from ((dummy variable for $T1^+ \ T2^- \ T3^-$) + β_2*Mean UE*). To obtain the negative incremental valuation consequences for missing the avoid loss threshold when all three thresholds have been made, we subtract ((dummy variable for $T1^+ \ T2^+ \ T3^+$) + β_8*Mean UE*) from ((dummy variable for $T1^- \ T2^+ \ T3^+$) + β_7*Mean UE*).

IV. MAIN RESULTS

Temporal Evidence on Threshold Hierarchy

Figure 1 shows temporal changes in the three managerial thresholds. Consistent with similar findings in Brown (2001), Bartov et al. (2002), Lopez and Rees (2002), Matsumoto (2002), and Dechow et al. (2003), we find the propensity to avoid negative surprises has increased over time (slope = −1.218, t-value = −13.91).[18] Consistent with Burgstahler and Eames (2003), we find the propensities to avoid losses (slope = −0.0361, t-value = −1.63) and earnings decreases are unchanged over time (slope = −0.1024, t-value = −1.62). A Chow test reveals that the slope of the propensity to avoid negative surprises exceeds that of the slope of the propensities to avoid either losses or earnings decreases at the 1 percent significance level or better.

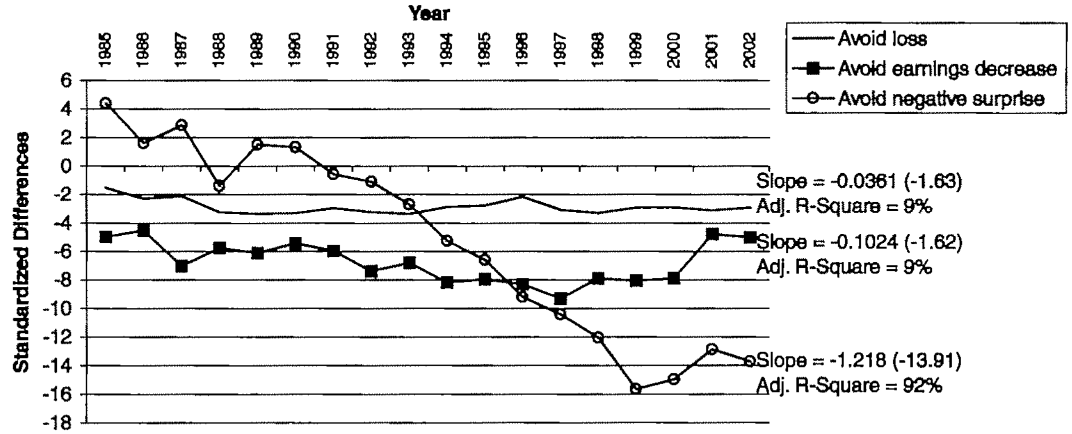
Consistent with Degeorge et al. (1999), managers seek to avoid losses or earnings decreases more than they do to avoid negative earnings surprises in each of the first nine years of our sample period, 1985–1993. However, in contrast to Degeorge et al. (1999), managers seek to avoid negative earnings surprises more than to avoid losses or earnings

---

[17] The 111,485 observations fall into the eight categories as follows (ordered as in the CAR equation with percent of total in parentheses): 8,367 (7.51 percent), 13,483 (12.09 percent), 1,923 (1.73 percent), 4,614 (4.14 percent), 14,697 (13.18 percent), 14,136 (12.68 percent), 3,080 (2.76 percent), and 51,185 (45.91 percent).

[18] Negative standardized differences indicate a propensity to achieve a threshold so a negative coefficient indicates an increased temporal propensity to do so. The interpretation of positive standardized differences in the early years is the well-known optimism documented in the literature for analyst forecasts in earlier time periods (O'Brien 1988; Das et al. 1998).

FIGURE 1
Temporal Analysis of Quarterly Earnings Thresholds



Temporal analysis of standardized differences. Yearly standardized differences are computed by first determining the distribution of deflated earnings levels, changes, and surprises in bins with .0025 widths. After setting sample size equal to 1000, standardized differences are computed as the difference between the actual and expected number of observations in a bin. Standardized differences for the bin just to the left of zero are plotted. A standardized difference with a negative value that has an absolute value of 2.33 or greater indicates threshold achievement. Slopes are computed using ordinary least squares regressions of standardized differences on year. For example, the slope of the avoid negative earnings surprise threshold is -1.218 with a t-statistic of -13.91 (p-value = 0.001), indicating that the propensity to avoid negative earnings surprises has increased significantly over time. This model provides 92 percent explanatory power.

decreases in each of the last seven years of our study, 1996–2002.[19] The results are strongly consistent with our first research question. Consistent with the findings of Dechow et al. (2003), but not with those of Graham et al. (2004), managers have a greater propensity to avoid negative quarterly earnings surprises in recent years than they do to avoid either losses or earnings decreases.

The Figure 1 evidence that avoiding earnings surprises is the most (least) sought-after threshold late (early) in our sample period suggests a temporal increase in management of analysts' quarterly earnings estimates.[20] Bartov et al. (2002) show that expectations management in the second half of their sample exceeded that in the first half of their sample.

---

[19] Beaver et al. (2003) argue that the discontinuity overestimates the actual amount of earnings management for avoiding both losses and earnings decreases. If such an overestimation pertains to our findings, the cross-over point may actually occur earlier in time than we suggest, but our major point remains: avoiding negative earnings surprises is the most sought-after threshold in recent years. An alternative explanation for our finding is that analyst forecasts have become timelier and thus less optimistic in recent years. To address this issue, we regress forecast timeliness (its proximity to earnings announcement date) on year. Finding no significant temporal trend, we reject this alternative interpretation of our results.

[20] Avoiding losses and avoiding earnings decreases can be achieved via earnings management, but not by expectations management. In contrast, avoiding negative earnings surprises can be accomplished both by earnings management and expectations management.

To add validity to their study and our Figure 1 findings, we use their methodology to calculate a proxy for the aggregate level of expectations management for each year in our sample period. For each year, we compute the difference in proportions of negative forecast error cases that end with a zero or positive surprise minus the proportion of positive or zero forecast error cases that end with a negative surprise (Bartov et al. 2002, H_7).[21] For simplicity, we do not provide a figure, but we correlate the Bartov et al. (2002) statistic for each year with absolute values of the three numbers in Figure 1 for the same year: *T3*, *T3−T2*, and *T3−T1*. The statistically significant Pearson correlations of 0.99, 0.97, and 0.98, respectively, provide strong evidence that expectations management explains why management of earnings surprises has increased *vis-à-vis* management of earnings levels and earnings changes.[22]

Temporal Evidence on the Valuation Consequences of Making Thresholds

Table 1 provides results of our yearly regressions. The intercept terms show the valuation consequences of missing (making) all three thresholds are negative (positive) every year, adding validity to our procedures. To see how avoiding negative surprises has become the threshold with the largest incremental valuation consequences in recent years, consider the coefficient estimates for our calculation of incremental valuation consequences for achieving one threshold (I_1, I_2, I_3, I_4, β_1, β_2, β_3, and β_4). If we only consider intercept and slope coefficients significant at the 5 percent (one-tailed) level or better in our calculations of incremental valuation consequences of making one threshold, then we can compare incremental valuation consequences of avoiding negative earnings surprises with avoiding losses and earnings decreases for the years 1994–2000. With the exception of 1995 in which there is a higher but insignificant coefficient for avoiding a negative earnings surprise, we find that for every other year, avoiding a negative earnings surprise unambiguously has the highest incremental valuation consequences. If we only consider significant intercept and slope coefficients when calculating the incremental valuation consequences for missing one threshold, then we find that during the 1993–2002 time period, not avoiding a negative earnings surprise has the most negative consequences and all of the avoid negative earnings

---

[21] Forecast error is defined as forecast revision plus earnings surprise. Consistent with Bartov et al. (2002), we define revision as the last analyst forecast for quarter *t* before quarter *t*'s earnings announcement minus the first analyst forecast for quarter *t* after the previous quarter's earnings announcement. Similar to our definition of the analyst forecast in our primary analysis, if there is more than one forecast on the day of the first forecast, then we take an average of the forecasts. To be consistent with Bartov et al. (2002), we require at least 20 trading days in between these two forecasts. Due to these additional data restrictions, our sample size is reduced to 70,530 observations. Although we use a different sample of firms due to our imposing additional constraints, we validate our calculation of Bartov et al.'s (2002) metric by computing proportions for the 1985–1993 and 1994–1997 subperiods. The proportion for 1985–1993 of 13.53 percent compares with 9.33 percent in Bartov et al. (2002) for their 1983–1993 subperiod. The proportion for 1994–1997 of 35.20 percent compares with 31.37 percent for their same 1994–1997 subperiod. The proportion for 1998–2002 (not included in Bartov et al. [2002]) is 50.98 percent.

[22] We also examined the relative frequency of negative surprises versus the relative frequency of negative forecast errors to see if the relative frequency of the latter exceeds the former, another test of expectations management (Bartov et al. 2002, H_6). We correlated this alternative Bartov et al. (2002) statistic for each year with absolute values of the three numbers in Figure 1 for the same year: *T3*, *T3−T2*, and *T3−T1*. The Pearson correlations of 0.86, 0.87, and 0.85 suggest that expectations management explains why management of earnings surprises has increased *vis-à-vis* management of earnings levels and earnings changes. Once again we validate our metric by comparing our subperiod results to Bartov et al.'s (2002) subperiod results. The proportion for 1985–1993 is 8.70 percent, which mirrors closely that found for Bartov et al.'s (2002) 1983–1993 subperiod of 8.48 percent. The proportion for 1994–1997 is 13.76 percent, which is nearly the same as 13.66 percent for their 1994–1997 subperiod. We further calculate the proportion for 1998–2002 (not included in Bartov et al. [2002]) to be 19.26 percent.

TABLE 1
Valuation Consequences of Achieving Thresholds

Year	n	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	Adj. R^2
1985	2,018	−0.0088***	−0.0064***	0.0024	−0.0092	−0.0026	0.0041	−0.0114	0.0106***	0.0046	0.1312	0.0005	0.6049	−0.3411	0.0813	0.5653	−0.0010	0.0260
1986	2,789	−0.0123***	−0.0118***	−0.0132**	−0.0045	0.0008	−0.0006	−0.0238	0.0165***	−0.0230	0.3439	0.0007	0.0215	0.0170	0.0767	0.7502	−0.0932	0.0584
1987	2,984	−0.0098**	−0.0106	−0.0083	0.0023	−0.0027	0.0044	0.0018	0.0099***	0.0206	1.0829**	−0.0190	−0.0392	0.0050	−0.0181	0.1828	0.0405	0.0279
1988	3,492	−0.0089***	−0.0115***	0.0105	−0.0009	−0.0075***	−0.0030	−0.0121	0.0047***	−0.0276***	0.1199***	0.4215***	−0.6079***	−0.6710***	0.3602**	−0.0077	0.0300	0.0412
1989	4,328	−0.0121***	−0.0095***	0.0070	0.0102	−0.0051***	0.0011	0.0134	0.0085***	0.0039	0.1964	0.2335	−0.8281**	0.0094	−0.0301	−0.0407	−0.0489	0.0305
1990	5,292	−0.0104***	−0.0124***	−0.0002	−0.0129	−0.0008	0.0043	0.0088	0.0164***	0.0084	−0.0498	0.2085***	0.4157**	−0.0622	0.0375	0.4123	0.0029	0.0391
1991	5,717	−0.0180***	−0.0126***	−0.0048	−0.0012	−0.0049**	0.0028	−0.0035	0.0172***	0.0181	0.1532	0.0486***	0.0590	−0.1959	1.1291***	0.0001	0.7993***	0.0683
1992	6,225	−0.0159***	−0.0175***	−0.0168***	−0.0158***	−0.0051***	0.0070**	0.0227***	0.0127***	0.1066*	0.2584	−0.2489**	0.1388	−0.0631	−0.3533	−0.1143	0.0477	0.0469
1993	6,415	−0.0210***	−0.0152***	0.0007	0.0075	−0.0067***	0.0011	−0.0019	0.0104***	−0.0055	0.8062***	0.0671	0.5411	−0.4943	1.6907***	−0.2136	1.4597***	0.0615
1994	7,240	−0.0227***	−0.0162***	−0.0229***	−0.0088**	−0.0116***	0.0020	0.0065	0.0096***	0.0042	1.3078***	−0.1591	−0.1177	−0.4681***	0.2762	−0.0046	1.3112***	0.0613
1995	8,678	−0.0176***	−0.0231***	−0.0115**	−0.0038	−0.0144***	0.0010	0.0047	0.0122***	0.0004	0.1851	−0.0540	0.4725	−1.0248**	0.1858	0.1322	0.9600***	0.0574
1996	8,878	−0.0238***	−0.0112***	0.0045	−0.0027	−0.0120***	−0.0005	0.0022	0.0136***	0.0827**	1.2599***	0.9416***	0.6104*	−0.2169	2.3409***	0.0208	1.2830***	0.0642
1997	9,503	−0.0293***	−0.0201***	−0.0135**	−0.0061*	−0.0114***	0.0005	0.0042	0.0131***	0.1087*	0.1440	0.4264**	0.0575	0.0625	2.3466***	−0.2122	0.5401***	0.0576
1998	9,337	−0.0292***	−0.0185***	−0.0109*	−0.0066**	−0.0120***	0.0040**	0.0146***	0.0100***	−0.0550	0.5374	0.0392	−0.0259	−0.7431	1.1503***	−0.2790	0.4656***	0.0340
1999	8,412	−0.0202***	−0.0190***	−0.0176*	−0.0037	−0.0138***	0.0071***	0.0047	0.0166***	0.1221***	0.0588	0.2608	−0.3035***	−0.6160	0.3572	−0.0562	0.0339	0.0365
2000	7,536	−0.0304***	−0.0209***	−0.0269***	−0.0099**	−0.0098**	0.0070***	−0.0064	0.0147***	0.0152	0.2452	−0.1534	0.0191	0.6106	0.1367	0.2686	1.2462***	0.0410
2001	7,248	−0.0265***	−0.0137***	−0.0088	0.0033	−0.0094*	0.0078***	0.0013	0.0169***	0.0197**	0.2474	−0.0795	−0.0462	−0.0242	1.2161***	−0.0014	0.2165	0.0348
2002	5,393	−0.0316***	−0.0189***	−0.0171**	−0.0060	−0.0184***	0.0052*	−0.0096**	0.0149***	−0.1857**	0.9624	0.1492	−0.1960	0.8000	1.7063***	0.0432	0.0245	0.0362

***,**,* Represents two-tailed significance levels of 1 percent, 5 percent, and 10 percent, respectively.
Coefficients reported above are from estimating the following regression:

$$CAR = \Sigma Category\ dummy + \beta_1{*}(T1^-\ T2^-\ T3^-){*}UE + \beta_2{*}(T1^+\ T2^-\ T3^-){*}UE + \beta_3{*}(T1^-\ T2^+\ T3^-){*}UE + \beta_4{*}(T1^-\ T2^-\ T3^+){*}UE + \beta_5{*}(T1^+\ T2^+\ T3^-){*}UE$$
$$+ \beta_6{*}(T1^+\ T2^-\ T3^+){*}UE + \beta_7{*}(T1^-\ T2^+\ T3^+){*}UE + \beta_8{*}(T1^+\ T2^+\ T3^+){*}UE + u. \tag{1}$$

T1, *T2*, and *T3* designate the first, second, and third thresholds (i.e., avoid losses, earnings decreases, and negative surprises), respectively, the superscripts minus and plus indicate if a threshold has been missed or made, *Category dummy* are categories of making/missing thresholds dummy variables and are denoted I_1–I_8, depending on the associated category, and *UE* is unexpected earnings, defined as actual earnings minus consensus analyst earnings forecast (deflated by beginning of quarter *t* stock price). For example, if a firm makes zero (all three) thresholds, the $T1^-\ T2^-\ T3^-$ ($T1^+\ T2^+\ T3^+$) dummy variables, I_1 and I_8, respectively, equal 1. The eight categorical dummy variables collectively represent the intercept of our model (i.e., the eight dummy vectors sum to the unit vector). Each of the eight dummies is interacted with the magnitude of unexpected earnings deflated by stock price, and are denoted by the slope coefficients, β_1–β_8 (earnings response coefficients). All coefficient estimates and adjusted R^2s are presented in decimal form.

surprise coefficients are significant, whereas the coefficients for the other two thresholds are insignificant for some years.[23]

Figure 2, Panel A presents a temporal analysis of the incremental valuation consequences of making one threshold. It is evident that before 1993, avoiding a negative surprise is not the obvious threshold to achieve first. Avoiding a negative surprise has the largest incremental valuation consequence in four of the first eight years of our sample period, 1985–1992. However, for all of the last ten years of our sample period (1993–2002), avoiding negative surprises has the largest incremental valuation consequence, making it clear that it is the threshold managers should seek to achieve most.

Figure 2, Panel B presents a temporal analysis of the incremental valuation consequences of missing one threshold. In the first four years of our sample period, 1985–1988, the valuation consequences of the threshold that is best to avoid missing is unclear. However, the valuation consequence of the threshold that is best to avoid missing is clear thereafter. In all of the last 14 years of our sample period, 1989–2002, reporting a negative surprise has the largest negative incremental valuation consequence.

The evidence is consistent with our second research question. It is clear that the positive (negative) valuation consequences of making (missing) thresholds are greatest for the third threshold in recent years but not in earlier years. Our results are consistent with the notion that managers shifted their threshold hierarchy (RQ_1) in response to changes in the incremental valuation implications of making/missing thresholds (RQ_2).

Shifts in Relative Accuracy and Precision of Analyst Forecasts

Before examining the temporal changes in the relative accuracy and precision of analyst forecasts (RQ_3 and RQ_4), we examine several other potential explanations for why investors may have shifted their focus in favor of avoiding negative earnings surprises. We examine if there was a temporal increase in the prevalence of analyst forecasts during our time period, and we summarize below (but do not tabulate) three types of related evidence.

(1) There is a temporal increase in the number of analysts in the I/B/E/S database over our time period (slope = 172.5, t-value = 14.19).
(2) There is a temporal increase in the number of firms in the I/B/E/S database over our time period (slope = 175.4, t-value = 5.75).
(3) There is a temporal increase in media attention (slope = 912.87, t-value = 12.69) paid to analyst forecasts over our time period.[24]

It is evident that the pervasiveness of analyst forecasts exploded during our time frame, increasing access to, and likely enhancing investor focus upon, analyst forecasts.[25]

To address our third research question, we examine whether analysts' earnings forecasts have become more accurate *vis-à-vis* a seasonal random walk model of quarterly earnings over our sample period. For this purpose, we conduct a temporal analysis of analyst forecast accuracy (the basis of our third threshold) *vis-à-vis* the accuracy of a seasonal time series model of quarterly earnings (the basis of our second threshold). Figure 3, Panel A plots

---

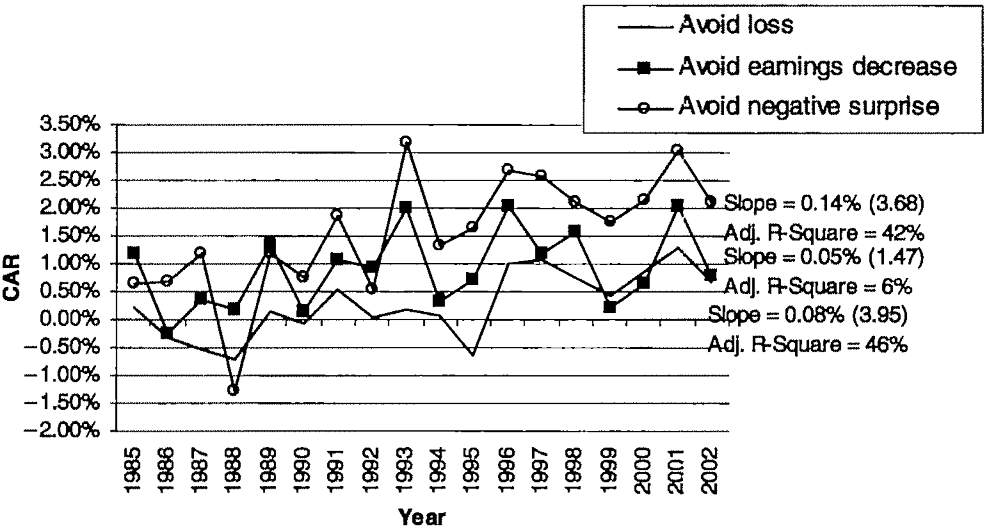
[23] We present all years in Figure 2 for presentation purposes, noting that we may be comparing insignificant coefficients for some years. However, as shown above, our main inference of the increased capital market importance of avoiding a negative earnings surprise is robust to considering only significant coefficients in the computation of incremental valuation consequences.

[24] We conducted a keyword search of analyst forecasts using the FACTIVA database. We performed keyword searches for analyst estimates, meet or beat earnings expectations, and consensus forecast and found similar results.

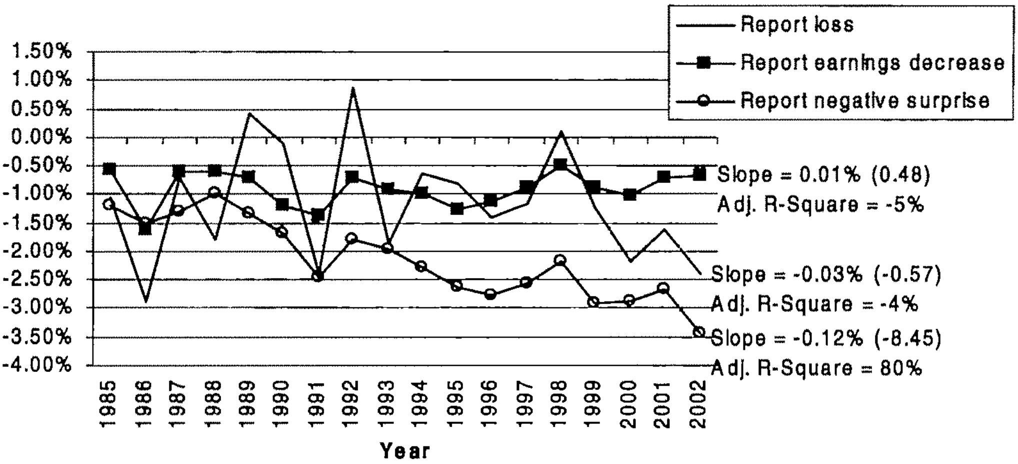
[25] The increased pervasiveness of analyst forecasts probably increased investor attention on analysts, while increased investor attention on analysts probably increased the pervasiveness of analyst forecasts.

FIGURE 2
Incremental Valuation Consequences of Quarterly Earnings Thresholds

Panel A: Making One Threshold



Panel B: Missing One Threshold



We estimate regression Equation (1) by year to obtain yearly coefficients for all eight mutually exclusive, collectively exhaustive categories of achieving thresholds. We then calculate the valuation consequence of making/missing a threshold by summing the value of the dummy variable for that threshold/year and the product of the appropriate interaction term coefficient and the mean level of deflated unexpected earnings, where *UE* is defined as actual earnings minus consensus analyst earnings forecast, for that threshold combination category for that year. We determine incremental valuation consequences by differencing the appropriate valuation consequences.

(*continued on next page*)

FIGURE 2 (continued)

Panel A presents a temporal analysis of the incremental (average) total valuation consequences of making one threshold. For example, for making the avoid loss threshold when no threshold has been made, we subtract ((dummy variable for $T1^- T2^- T3^-$) + β_1*Mean *UE*) from ((dummy variable for $T1^+ T2^- T3^-$) + β_2*Mean *UE*).

Panel B presents a temporal analysis of the incremental (average) total valuation consequences of missing one threshold. For example, for missing the avoid loss threshold when all three thresholds have been made, we subtract ((dummy variable for $T1^+ T2^+ T3^+$) + β_8*Mean *UE*) from ((dummy variable for $T1^- T2^+ T3^+$) + β_7*Mean *UE*).

We compute the slopes using ordinary least squares regressions of valuation consequences on year. For example, in Panel A, the slope of the avoid earnings decrease threshold is 0.05 percent with a t-statistic of 1.47. The model which regresses the avoid earnings decrease incremental valuation consequences on the integers 1 to 18 for years = 1985 to 2002 provides 6 percent explanatory power.

the median of the distribution of absolute values of the ratio of analysts' earnings surprises to earnings changes.[26] The adjusted R^2 of the regression of this ratio on time is 92 percent, and the t-value on the time variable is −14.22, providing strong evidence of a temporal increase in the relative accuracy of analyst forecasts *vis-à-vis* that of a seasonal random walk time-series model.

To address our fourth research question, we examine whether the precision of analysts' quarterly earnings forecasts has increased during our sample period. Figure 3, Panel B plots the median forecast dispersion over time.[27] The adjusted R^2 is 87 percent and the t-value on the time variable is −10.64, providing strong evidence of a temporal increase in the precision of analysts' quarterly earnings forecasts. Our Figure 3 findings are consistent with our third and fourth research questions.

V. SENSITIVITY ANALYSES

We now address five potential validity threats: (1) use of the procedure suggested by Burgstahler and Dichev (1997); (2) use of split-adjusted data, (3) changing composition of sample firms during our time period; (4) using I/B/E/S actuals to measure earnings levels and changes; and (5) using non-GAAP earnings. For simplicity, we do not tabulate results using procedures that mitigate these validity threats, but we summarize our findings below.

Ratio Statistic

In our primary analysis, we computed a test statistic setting sample sizes equal to 1,000 (see footnote 15) to avoid using test statistics that are sensitive to sample size. Because this procedure ignores actual sample sizes, we employ a nonparametric statistic used by Dechow et al. (2003), whose magnitude is unaffected by sample size. We compute the ratio as the actual number of observations in the bin just to the right of (and including) zero divided by the actual number of observations in the bin just to the left of zero. We obtain qualitatively similar results as in Figure 1 except that the avoid negative surprise threshold becomes the most sought after threshold two years later (in 1998). Similar to Figure 1,

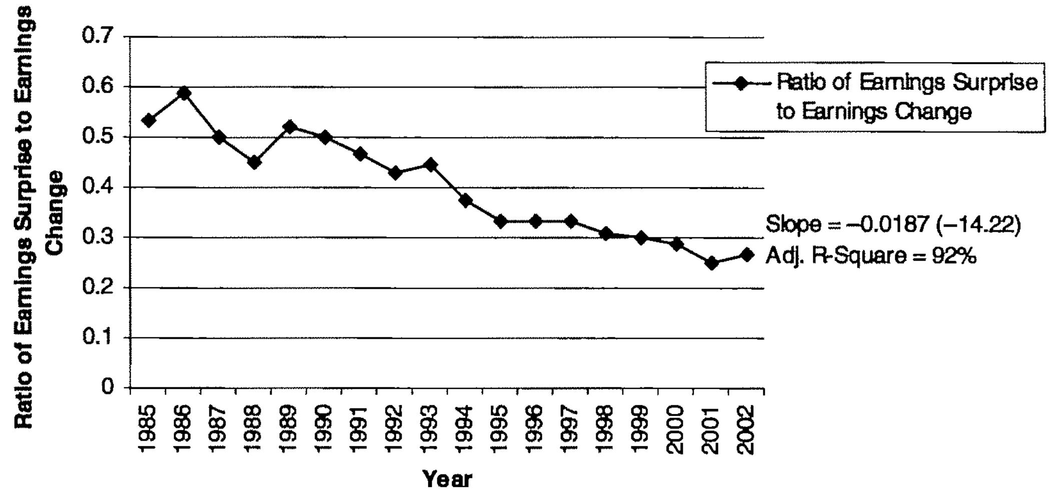
---

[26] For this analysis, we delete any observation where the earnings surprise or earnings change is equal to 0.

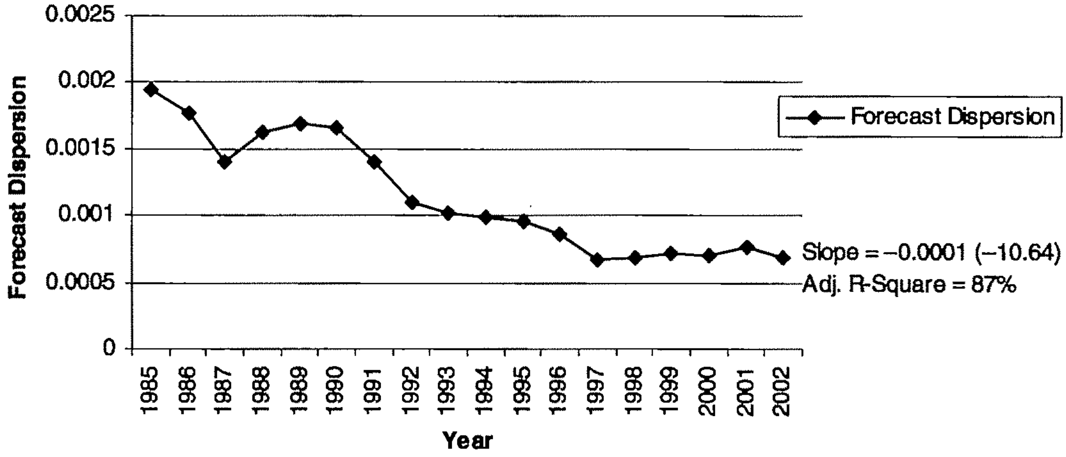
[27] To ensure a reasonable number of observations are available to measure dispersion, we require that at least three analysts follow the firm.

FIGURE 3
Relative Accuracy and Forecast Dispersion over Time

Panel A: Relative Accuracy of the Forecast Metrics over Time



Panel B: Analyst Forecast Dispersion over Time



Panel A presents a temporal analysis of the median ratio of earnings surprises to earnings changes. Sample size is 93,829 observations. For each year we compute the median of the distribution of absolute values of the ratio of earnings surprises to earnings changes for all firm-quarters in a particular year.

Panel B presents a temporal analysis of forecast dispersion. Sample size is 82,473. For each year we compute the median forecast dispersion for all firm-quarters in a particular year. Consistent with Kinney et al. (2002), we define dispersion of analyst forecasts as the standard deviation of forecasts as of the last consensus forecast update before the quarterly earnings announcement, scaled by price per share. We also require that at least three individual analysts make a forecast at that date. We compute the slopes using ordinary least squares regressions of the median ratio of earnings surprises to earnings changes or the median forecast dispersion on year. We report the slope (t-value) and adjusted R^2 of these regressions.

avoiding negative earnings surprises is the only threshold with a significant temporal increase (slope = 0.251, t-value = 11.32).

Split-Adjusted Data

Baber and Kang (2002) show that errors induced by rounding split-adjusted data may alter conclusions about how investors react to earnings that exactly meet analyst expectations. We mitigate this potential validity threat in two different ways. First, we delete all observations with earnings levels, earnings changes, or earnings forecast errors equal to zero cents per share, reducing our sample size to 93,273. Second, we use the I/B/E/S procedure and file (with adjustment factors) for un-adjusting for the stock splits. Both our threshold and valuation consequence results are qualitatively unchanged using either procedure, indicating that our results are not due to our failure to adequately account for stock splits.

Same Firms in Sample during Entire Time Period

Our results may arise from a shift in composition of sample firms during our time period. To mitigate this potential validity threat, we restrict our analyses to firms that are in our sample during our entire time period, reducing our sample size to 24,016. Our threshold and valuation consequence results are qualitatively similar when we limit our sample to these firms, suggesting that our results are not due to temporal shifts in sample composition.

Compustat versus I/B/E/S Actuals

I/B/E/S is primarily a source of earnings forecast data, not a vendor of actual earnings numbers so there is likely to be less measurement error in computing the third threshold, especially in recent years when the analyst data providing business has become so competitive, when First Call came on the scene in the early 1990s.[28] To examine the seriousness of this validity threat, we re-compute the first two thresholds using quarterly Compustat earnings data (Compustat data item 9 or 19, depending on whether the I/B/E/S forecast data is reported on a diluted or primary (basic) basis, respectively) in lieu of I/B/E/S earnings data, deflated by Compustat stock price, reducing our sample size to 89,628.[29] We obtain qualitatively similar threshold avoidance and valuation consequence results using Compustat data to compute our first two thresholds, revealing that our results are not attributable to use of I/B/E/S actuals to compute our first two thresholds.

Pro Forma versus GAAP Earnings

The change in relative accuracy we document may result from analysts predicting proforma earnings rather than GAAP earnings more over time (Bradshaw and Sloan 2002; Bhattacharya et al. 2004). Use of non-GAAP earnings (i.e., I/B/E/S actuals) may "stack the deck" in favor of analyst forecasts and actuals (i.e., our third threshold), especially in recent years when the use of pro forma has become so pervasive. To mitigate this validity threat, we examine the relative accuracy of analyst forecasts *vis-à-vis* that of a seasonal random walk model over time for the subset of firms where I/B/E/S actuals equal S&P

---

[28] Philbrick and Ricks (1991) compare Value Line and I/B/E/S actuals with Compustat earnings numbers. Finding more concordance of the Value Line actuals with the Compustat actuals, they conclude that Value Line provides better actuals. Ramnath et al. (2005) show that Value Line does not provide better actuals than I/B/E/S in recent years.

[29] We continue to use I/B/E/S data to compute our third threshold because we do not wish to mix actuals and forecasts from different sources (Philbrick and Ricks 1991).

GAAP earnings (Compustat data item 9 or 19, depending on whether the I/B/E/S forecast data is reported on a diluted or primary (basic) basis, respectively). We find a significant temporal increase in relative accuracy even for this subset of firms. The adjusted R^2 of the regression of this ratio on time is 72 percent and the t-value on the time variable is –6.66. It is evident that our results are not driven by use of pro forma earnings numbers in lieu of GAAP earnings.

VI. CONCLUSIONS AND IMPLICATIONS

Degeorge et al. (1999) use pooled temporal, cross-sectional data, and conclude that managers seek to avoid quarterly losses and earnings decreases more than to avoid negative earnings surprises. A 2003 survey by Graham et al. (2004) finds that managers attest that they seek to avoid quarterly earnings decreases more than to avoid negative quarterly earnings surprises. In contrast to these survey results, Dechow et al. (2003) apply a Degeorge et al. (1999) type methodology to annual data, and find that managers are more prone to avoid negative annual earnings surprises than to avoid either losses or earnings decreases in the last three years of their study, 1999–2001. These conflicting results lead us to re-examine threshold hierarchies applying a Degeorge et al. (1999) type methodology to quarterly data.

We confirm Degeorge et al.'s (1999) results earlier in our sample period, but we find precisely the opposite later in our sample period. Our findings reveal that the quarterly hierarchy proposed by Degeorge et al. (1999) is no longer valid, and that the threshold managers maintain they seek to achieve most (Graham et al. 2004) is inconsistent with the threshold that they actually seek to achieve most. We appeal to economic intuition to examine why managers changed their threshold prioritization. We assume that managers are wealth-maximizers who recognize that their wealth (i.e., stocks, options, and human capital) is adversely impacted when their firms' stock prices drop if they fail to achieve thresholds. Our finding that capital markets unambiguously reward (sanction) firms more for making (missing) analyst forecasts late, but not early, in our time period, is consistent with the notion that managers took a cue from the capital markets to alter their threshold prioritization in favor of avoiding negative quarterly earnings surprises.

We examine why investors changed their focus in favor of analyst forecasts. We find a significant increase over our sample period in the number of analysts, the number of firms followed by analysts, the amount of media attention paid to analyst forecasts, the relative accuracy of analyst forecasts *vis-à-vis* that of a seasonal random walk model, and the precision of analyst forecasts. All of these factors may have induced investors to pay more attention to analysts' forecasts later than earlier in our time period.

Our findings are robust to five sensitivity checks: (1) using a ratio metric suggested by Dechow et al. (2003) in lieu of a Burgstahler and Dichev (1997) measure; (2) un-adjusting earnings data for stock splits; (3) requiring firms to be in the sample for the entire time period; (4) using Compustat actuals in lieu of I/B/E/S actuals; and (5) omitting observations where I/B/E/S earnings differ from GAAP earnings. These results increase our confidence in our findings of temporal shifts in managers' threshold hierarchy and the incremental valuation consequences thereof that favor avoidance of negative quarterly earnings surprises.

We close with two caveats. First, our findings are subject to the validity of our methodologies. It is possible that the survey results of Graham et al. (2004) are more valid than the Burgstahler and Dichev (1997)/Degeorge et al. (1999)/Dechow et al. (2003) type methodologies to examine thresholds. However, it is at least as plausible that managers report in surveys that they focus on one thing (avoid quarterly earnings decreases) while they actually focus more on something else (avoid quarterly negative surprises). Managers may do this unwittingly or they may report something in a survey that they believe the surveyor wants them to state. Second, we do not definitively show that managers altered their threshold prioritization *because* investors changed the thresholds they reward (sanction) the most. Similarly, we do not definitively show that investors altered the threshold they reward (sanction) most for making (missing) *because* of the proliferation of analyst forecasts, media attention paid to analysts, and/or the increase in the accuracy and precision of analyst forecasts.

REFERENCES

Baber, W., and S. Kang. 2002. The impact of split adjusting and rounding on analysts' forecast error calculations. *Accounting Horizons* 16 (December): 277–289.

Barber, B., and J. Lyon. 1999. Improved methods for tests of long-run abnormal stock returns. *Journal of Finance* 54 (February): 165–201.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (June): 173–204.

Beaver, W., M. McNichols, and K. Nelson. 2003. An alternative interpretation of the discontinuity in earnings distributions. Working paper, Stanford University, Palo Alto, CA.

Bhattacharya, N., E. Black, T. Christensen, and R. Mergenthaler. 2004. Empirical evidence on recent trends in pro forma reporting. *Accounting Horizons* 18 (March): 27–43.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the Street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (March): 41–66.

Brown, L., P. Griffin, R. Hagerman, and M. Zmijewski. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9 (April): 61–87.

———, and K. Kim. 1991. Timely aggregate analyst forecasts as better proxies for market earnings expectations. *Journal of Accounting Research* 39 (Autumn): 382–385.

———. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (September): 221–241.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (December): 99–126.

———, and M. Eames. 2003. Earnings management to avoid losses and small decreases: Are analysts' fooled? *Contemporary Accounting Research* 20 (Summer): 253–294.

Das, S., C. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73 (April): 277–294.

Dechow, P., S. Richardson, and I. Tuna. 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* 8 (June–September): 355–384.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (January): 1–33.

Eccles, R. G., R. H. Herz, E. M. Keegan, and D. M. H. Phillips. 2001. *The Value Reporting Revolution: Moving Beyond the Earnings Game*. New York, NY: John Wiley & Sons, Inc.

Fried, D., and D. Givoly. 1982. Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics* 4 (October): 85–107.

Graham, J., C. Harvey, and S. Rajgopal. 2004. The economic implications of corporate financial reporting. Working paper, Duke University, Durham, NC.

Imhoff, E., and G. Lobo. 1992. The effect of ex-ante earnings uncertainty on earnings response coefficients. *The Accounting Review* 67 (April): 427–439.

Kinney, W., D. Burgstahler, and R. Martin. 2002. Earnings surprise "materiality" as measured by stock returns. *Journal of Accounting Research* 40 (December): 1297–1329.

Libby, R., and W. Kinney. 2000. Does mandated audit communications reduce opportunistic corrections to manage earnings to forecasts? *The Accounting Review* 75 (October): 383–404.

Lopez, T., and L. Rees. 2002. The effect of beating and missing analysts' forecasts on the information content of unexpected earnings. *Journal of Accounting, Auditing, and Finance* 17 (Spring): 155–184.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (July): 483–514.

Ng, T., and H. Tan. 2003. Why do auditors waive quantitatively immaterial audit differences that affect clients' ability to meet analysts' expectations? Working paper, Nanyang Technological University, Nanyang, Singapore.

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10 (January): 53–83.

Philbrick, D. R., and W. E. Ricks. 1991. Using Value Line and I/B/E/S forecasts in accounting research. *Journal of Accounting Research* 29 (Autumn): 397–417.

Ramnath, S., S. Rock, and P. Shane. 2005. Value Line and I/B/E/S earnings forecasts. *International Journal of Forecasting* 21 (January-March): 185–198.

Soffer, L., S. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5 (March): 5–26.

Trieschmann, J., A. Dennis, G. Northcraft, and A. Niemi. 2000. Serving multiple constituencies in business schools: MBA program versus research performance. *Academy of Management Journal* 43 (December): 1130–1141.

Wiedman, C. 1996. The relevance of characteristics of the information environment in the selection of a proxy for the market's expectations for earnings: An extension of Brown, Richardson and Schwager [1987]. *Journal of Accounting Research* 34 (Autumn): 313–324.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 441–476

Equity Incentives and Earnings Management

Qiang Cheng
University of British Columbia

Terry D. Warfield
University of Wisconsin—Madison

ABSTRACT: This paper examines the link between managers' equity incentives—arising from stock-based compensation and stock ownership—and earnings management. We hypothesize that managers with high equity incentives are more likely to sell shares in the future and this motivates these managers to engage in earnings management to increase the value of the shares to be sold. Using stock-based compensation and stock ownership data over the 1993–2000 time period, we document that managers with high equity incentives sell more shares in subsequent periods. As expected, we find that managers with high equity incentives are more likely to report earnings that meet or just beat analysts' forecasts. We also find that managers with consistently high equity incentives are less likely to report large positive earnings surprises. This finding is consistent with the wealth of these managers being more sensitive to future stock performance, which leads to increased reserving of current earnings to avoid future earnings disappointments. Collectively, our results indicate that equity incentives lead to incentives for earnings management.

Key words: *equity incentives; stock-based compensation; stock ownership; insider trading; earnings management; meeting or beating analysts' forecasts; abnormal accruals.*

Data Availability: *Data used in this study are available from the sources identified in the text.*

We thank anonymous referees, Holly Ashbaugh, Joy Begley, Sandra Chamberlain, Changling Chen, Jerry Feltham, Steve Henning, Mark Huson, Bin Ke, Michael Kimbrough, Mark Kohlbeck, Sung S. Kwon, Ryan LaFond, Matt Magilke, Brian Mayhew, Per Olsson, Jamie Pratt, Katherine Schipper, Terry Shevlin, D. Shores, Jim Wahlen, John Wild, workshop participants at Indiana University and University of Wisconsin–Madison, conference participants at the 2004 Financial Accounting and Reporting Section Midyear Conference in Austin, TX, and the 2004 Canadian Academic Accounting Association Annual Meeting in Vancouver, Canada, especially Xia Chen and Patty Dechow for helpful comments on earlier versions of this paper. Professor Warfield gratefully acknowledges the financial support of the PricewaterhouseCoopers Foundation.

Editor's note: This paper was accepted by Patricia Dechow, Editor.

Submitted October 2003
Accepted July 2004

I. INTRODUCTION

Regulators and investors have raised concerns that certain management incentives could lead to earnings management, reducing the informativeness of financial reporting and contributing to recent corporate scandals (Levitt 1998; Knowledge at Wharton 2003). Of particular recent concern is that stock-based compensation might induce managers to increase the short-term stock price through earnings management. For example, Michael Jensen argues that stock-based compensation and other equity incentive elements can encourage managers to increase short-term stock prices in order to benefit from subsequently selling shares of their own firms' stock (*The Economist* 2002).[1] Indeed, these concerns have contributed to new calls for the expensing of stock options, other regulatory constraints on compensation practices, and improved corporate governance (FASB 2004; Morgenson 2003; Conference Board 2002). However, there is little empirical evidence on the relation between equity incentives and earnings management. The purpose of this paper is to examine this relation, thereby providing evidence that can be used to corroborate or refute allegations of equity-incentive-related earnings management.

Modern corporations employ various mechanisms to remedy the adverse consequences arising from the separation of ownership and control (Jensen and Meckling 1976). Encouraging managerial ownership is an example of such mechanisms. Theoretically, as managers own more shares, they are more likely to act in the interests of shareholders. One way to increase the level of managerial ownership is to award managers options or shares of stock (Core and Guay 1999). Prior research examines the association of managerial ownership and stock-based compensation with future firm performance, and finds evidence consistent with the incentive-alignment effect of these equity incentive elements (Lambert and Larcker 1987; Morck et al. 1988; Hanlon et al. 2003, among others). However, at the same time, these elements can lead managers to focus on short-term stock prices, thereby leading to incentives for earnings management.

The incentives for earnings management arise from managers' diversifying the increased risk associated with ownership or stock-based compensation. For example, Ofek and Yermack (2000) examine the dynamics of ownership and stock-based compensation, and find that when managers are awarded stock-based compensation, they tend to sell shares they already own for risk diversification reasons. Managers are likely to continue selling shares in the future if the risk exposure due to equity incentives is, or is expected to be, above the level that managers are willing to bear. The risk exposure can increase in the future because when stock prices increase, managerial wealth becomes more concentrated in the stock and managers' option holdings become riskier (i.e., more sensitive to stock prices).

Because of this selling, a manager's wealth is sensitive to the short-term stock price, which can motivate managers with high equity incentives to increase the short-term stock price. Given that the capital market uses current earnings to predict future earnings when pricing firm equity, these managers are expected to use their accounting discretion to manage earnings in order to keep the short-term stock price high (Stein 1989).

In this paper, we explicitly examine the links from equity incentives to earnings management. In particular, we examine the following equity incentive elements: option grants, unexercisable options, exercisable options, restricted stock grants, and stock ownership. Because the positive relation between equity incentives and managers' future sales of their own firms' stock underpins the earnings management argument, we first document the existence of this relation. We measure a manager's future net sales as the difference between open market sales and the sum of open market purchases and the shares received from exercising options. The results indicate that, as expected, managers with high equity incentives have significantly higher levels of net sales in the year after earnings announcements.[2]

---

[1] Dechow and Skinner (2000) also express concerns that when stock market valuations are high, managers have incentives to maintain or even increase those valuations, probably via earnings management.

Next, we investigate the association between equity incentives and earnings management by examining the extent to which managers report earnings that meet or just beat analysts' forecasts. As expected, we find a significantly higher incidence of meeting or just beating analysts' forecasts for firms with higher managerial equity incentives. Based on analyses that control for firm performance and other potential confounds, we conclude that our results are more consistent with earnings management induced by equity incentives as opposed to better firm performance attendant to incentive alignment benefits or to equity incentive elements correlating with other firm characteristics.

We then examine whether managers who predictably engage in earnings management sell more shares after actual earnings management. We find that managers with high equity incentives sell more after meeting or beating analysts' forecasts than after missing analysts' forecasts. In contrast, we do not find such evidence for managers with low equity incentives. These results are consistent with high equity incentive managers increasing sales after earnings management.

We also conduct additional analyses to enrich our results on earnings management. First, equity incentives can be recurring; managers with high equity incentives in the current period are more likely to have high equity incentives in the future. Thus, managers with high equity incentives can benefit from smoothing earnings. By reserving in good years, these managers can more easily meet or beat analysts' forecasts in future periods. Consistent with this argument, we find that high equity incentive managers are less likely to report large positive earnings surprises compared to those with low equity incentives. The evidence is particularly strong for firms with more persistent equity incentives. Second, we find that high equity incentive managers, on average, report more income-increasing abnormal accruals, especially for managers with less persistent equity incentives—those who are less concerned with the reversal of accruals. We also find that managers sell more shares after taking income-increasing abnormal accruals. Last, our findings are robust to alternative measurements of key variables, the inclusion of other control variables, and alternative regression specifications.

Overall, consistent with regulators' and investors' concerns about equity incentives, we find that high equity incentive managers are more likely to engage in earnings management, relative to managers with low equity incentives. This evidence contributes to the compensation literature by extending research on compensation-related opportunistic behavior. By documenting one potential consequence of stock-based compensation and ownership, this paper should be of interest to boards of directors contemplating compensation contracts for managers. While stock-based compensation and ownership can yield positive incentive alignment effects, optimal compensation contracts should reflect the earnings management consequences arising from these equity incentives.

This paper also contributes to the earnings management literature. Although earnings management has received considerable attention in the accounting literature, little is known about the earnings management incentives arising from stock-based compensation or ownership. This paper therefore extends research on compensation-related earnings management beyond that related to bonus contracts as documented in Healy (1985) and Warfield et al. (1995).

---

[2] In our main analyses, we focus on CEOs because they are the ultimate decision makers. Results based on all managers with data available are similar.

Our paper is related to other concurrent studies of the capital market pressure on managers due to equity incentives. Ke (2001) finds that managers' equity incentives are positively correlated with the duration of a string of quarterly earnings increases and Skinner and Myers (2000) find that firms tend to manage earnings to maintain a string of earnings increases. Beneish and Vargus (2002) find that income-increasing accruals, when accompanied with insider sales, are of lower quality, consistent with the notion that managers might manage earnings upward and then sell their own shares. We extend these studies by explicitly linking equity incentives to financial reporting attributes that reflect earnings management (meeting or just beating analysts' forecasts, avoiding large positive earnings surprises, and recognizing abnormal accruals) and by providing direct evidence of managers trading in response to such earnings management.[3]

The remainder of the paper is organized as follows. In the next section, we develop our hypothesis concerning equity incentives and earning management. In Section III, we describe the sample and data with primary results presented in Section IV. Section V discusses further analyses that provide additional insights and ensure the robustness of our results; Section VI concludes with a summary.

II. HYPOTHESIS DEVELOPMENT

The theoretical linkages for our study are illustrated in Figure 1. As indicated, the link between equity incentives and managers' future trading is important because such trading motivates managers to care about short-term stock prices, resulting in earnings management incentives. In this section, we first discuss the relation between equity incentives and managers' future trading of their own firms' stock, and then present our hypothesis on earnings management.

Equity Incentives and Their Relation to Future Manager Trading

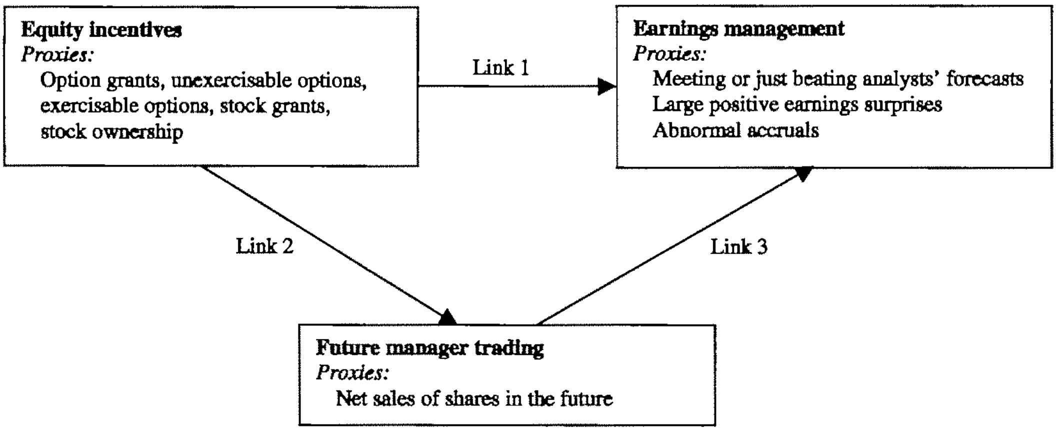
Managerial ownership is an important mechanism to align managers' incentives with those of shareholders (Jensen and Meckling 1976; Morck et al. 1988). One way to increase ownership is through stock-based compensation. When a manager is granted options, the options usually are not exercisable until three or four years later. When options become exercisable (i.e., vested), managers can choose to hold exercisable options or to exercise the options and hold shares instead. (To finance option exercises, managers generally sell the shares received from exercising options.) Managers are also awarded restricted stock, which, like option grants, does not vest (i.e., cannot be traded) until after three or four years. Another way to obtain shares is through open market purchases, which is common in early career stages (Core and Larcker 2002).

These various elements and their interrelationships are summarized in Figure 2. As indicated, at any point in time, such as at the fiscal-year-end, managers might have various forms of equity-based holdings: unexercisable options, exercisable options, and stock ownership. Due to these equity-based holdings, managers' wealth is sensitive to their firms' stock prices and managers therefore bear the idiosyncratic risk of the firm.

---

[3] Concurrent research by McVay et al. (2004) examines insider trading *following* meeting and beating analysts' forecasts. Unlike this paper, McVay et al. (2004) do not examine *ex ante* incentives for earnings management resulting from stock-based compensation or ownership.

FIGURE 1
Theoretical Links


Equity incentives
Proxies:
Option grants, unexercisable options, exercisable options, stock grants, stock ownership
Link 1
Earnings management
Proxies:
Meeting or just beating analysts' forecasts
Large positive earnings surprises
Abnormal accruals
Link 2
Link 3
Future manager trading
Proxies:
Net sales of shares in the future

This figure depicts the theoretical constructs and the links between them. For each theoretical construct, we list empirical proxies used in the analyses. Tests of Link 1 comprise our main analyses, in which we investigate whether high equity incentive managers are more likely to manage earnings. While we primarily rely on prior research to establish Link 3, we test Link 2 to verify that high equity incentive managers sell more in the future.

The equilibrium risk level results from a balance of two opposing forces. From an incentive-alignment perspective, equity incentives help motivate managers to work in shareholders' interests, thereby reducing agency costs. However, from the perspective of risk diversification, risk-averse managers want to reduce their exposure to the idiosyncratic risk of the firm. Consequently, managers want to sell shares if the risk exposure level is higher than the equilibrium level.[4] Even if the current risk exposure is at the equilibrium level, it can increase beyond this level in the future due to changes in economic circumstances.

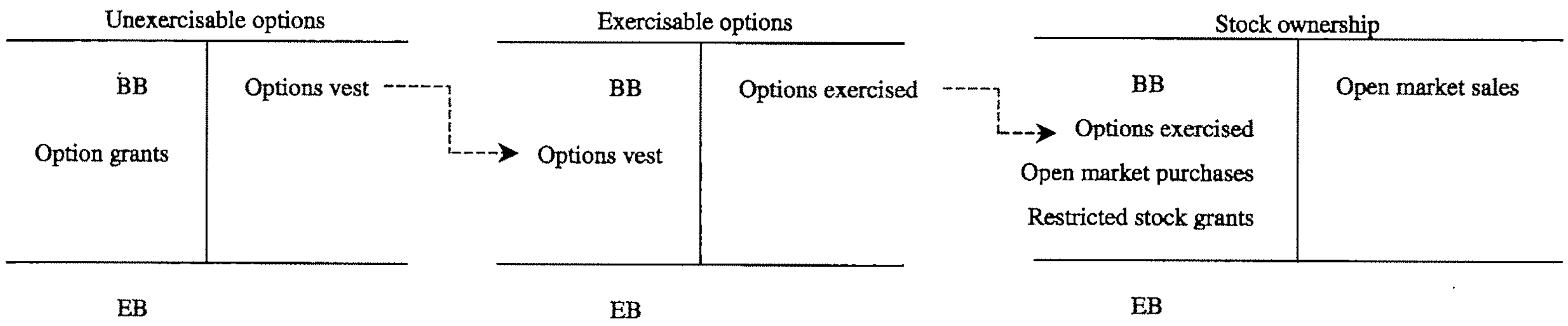
For example, if stock prices increase, then managers' wealth becomes more concentrated in a single risky asset (i.e., their firms' stock). Thus, the exposure of managers' wealth to the idiosyncratic risk increases. Furthermore, based on option valuation theory (such as Black and Scholes 1973), the sensitivity of the value of unexercisable options to stock price (i.e., the hedge ratio or the delta of the option) increases with stock price.[5] Under general capital asset pricing assumptions, the expected return of a firm's stock is positive; that is, stock prices are expected to increase in normal economic circumstances. As a result, managers' risk exposure due to equity incentives predictably increases in the future. If it increases beyond the equilibrium level, then managers might sell shares they already own to diversify the risk.

---

[4] If managers' equity incentive levels are lower than the equilibrium level, then they can purchase shares from the open market. Some firms explicitly specify the target stock ownership for CEOs to facilitate incentive alignment (Core and Larcker 2002).

[5] One can formally derive this result from examining the Black and Scholes (1973) formula. The intuition is as follows. The hedge ratio of an option equals the probability that the option is in the money when it is exercisable. (The hedge ratio of an in-the-money exercisable option is 1.) Since the probability for an option to be in the money when it is exercisable increases with stock prices, its hedge ratio increases with stock prices.

FIGURE 2
The Dynamics of Equity Incentives



This figure uses T-accounts to depict the relation between various components of equity incentives, and the relation between equity incentives and manager trading (i.e., open market purchases, options exercised, and open market sales). BB denotes beginning balance and EB denotes ending balance.

In summary, managers with higher equity incentives are more likely to sell shares (Link 2 in Figure 1) and their wealth is therefore sensitive to short-term stock prices. This sensitivity can lead to incentives to increase short-term prices and the value of the shares managers are going to sell. To this end, they can make value-increasing production, financing, and investment decisions, and/or engage in earnings management. The latter is the focus of this study.

Equity Incentives and Earnings Management

Managers who sell shares in the future have earnings management incentives if two conditions hold: (1) the capital markets rely on reported earnings in forming beliefs about future earnings so that earnings management can affect stock prices, and (2) managers can take advantage of the increased stock prices. With respect to the first condition, Stein (1989, 657) argues that "the stock market uses earnings to make a rational forecast of firm value—higher earnings today will be correlated with higher earnings in the future." Extant accounting research shows that stock prices generally react to earnings announcements in the direction of earnings surprises, with analysts' forecasts as the proxy for earnings expectations (O'Brien 1988; Brown and Kim 1991; among others). Thus, managers can use their accounting discretion to affect reported earnings and stock prices if the capital markets have difficulty in detecting earnings management.[6]

If stock prices increase with earnings management, then managers gain more from selling shares in the future than without earnings management. Thus, condition (2) is satisfied, unless litigation or reputation concerns deter managers from taking advantage of the increased stock price. The evidence in prior research is consistent with managers being able to benefit from insider trading. For example, Seyhun (1998) and Lakonishok and Lee (2001), among others, find that insiders are generally better informed and earn abnormal returns. With respect to accounting information disclosed through management forecasts or earnings announcements, Penman (1982), Sivakumar and Waymire (1994), and Noe (1999) find that managers sell more after good news than after bad news. Summers and Sweeney (1998) and Beneish (1999) find that managers of firms accused of accounting fraud sell their own shares before these firms are formally subject to SEC enforcement actions. Similarly, Beneish and Vargus (2002) find evidence consistent with managers selling more after market participants overprice income-increasing accruals.

Thus, extant research supports the link from insider trading to earnings management (Link 3 in Figure 1). This link, combined with the above-discussed link from equity incentives to future trading, suggests that high equity incentive managers are more likely to engage in earnings management, relative to low equity incentive managers (Link 1 in Figure 1). This prediction is consistent with the conclusion in Stein (1989): managers who care about short-term stock prices are expected to manage earnings to increase short-term stock prices and such behavior increases with the sensitivity of managers' utility to current stock prices.[7]

---

[6] While the efficient market assumption posits that managers cannot fool investors with earnings management, extant empirical evidence suggests otherwise. For example, Kasznik (1999) finds that earnings management can help companies avoid market penalties associated with missing earnings targets. Sloan (1996) and Xie (2001), among others, suggest that the capital markets misprice accruals, especially abnormal accruals, consistent with earnings management influencing stock prices. See Fields et al. (2002) for additional discussion of the pricing effects of earnings management. As discussed below, incentives for earnings management exist in our setting even if on average earnings management cannot affect stock prices.

Last, earnings management incentives exist not only when investors are naïve—fixating on reported earnings, but also when investors are sophisticated—adjusting for expected earnings management when pricing equity. Stein (1989) shows that when investors are sophisticated, in equilibrium, investors rationally expect managers to engage in earnings management, which they do. In other words, if the market has rational expectations of earnings management behavior but cannot completely undo earnings management of individual firms, it is optimal for market participants to price the reported earnings by adjusting for the impact of earnings management. Given investors' pricing behavior, managers find it optimal to manage earnings.

In sum, as depicted in Figure 1, managers with high equity incentives are likely to sell shares. Such trading behavior motivates managers to care about short-term stock prices and to manage earnings. Formally, we hypothesize that:

> **H1:** The incidence of earnings management is positively associated with equity incentives.

In this paper, we focus on one particular proxy for earnings management—meeting or *just* beating analysts' forecasts. Recent studies of earnings management suggest that the disproportionate likelihood of meeting or *just* beating analysts' forecasts is an important manifestation of earnings management (Degeorge et al. 1999; Burgstahler and Eames 2003; among others). We focus on this proxy because it has more direct market consequences relative to other proxies, such as abnormal accruals. Prior research has documented the reward for meeting analysts' forecasts (Bartov et al. 2002; Kasznik and McNichols 2002) and the negative market consequences of missing analysts' forecasts (Skinner and Sloan 2002). To further corroborate and enrich our earnings management results, in Section V we also analyze the likelihood of large positive earnings surprises to investigate the existence of earnings smoothing and the recognition of abnormal accruals.

III. SAMPLE AND DATA

The initial sample includes all firm-years with data on CEOs' stock-based compensation and ownership available from the Standard & Poor's ExecuComp database for the period 1993–2000.[8] We exclude financial institutions (SICs between 6000 and 6999) and utilities (SICs between 4400 and 5000) because managers in these regulated industries might have different motivations to manage earnings (Burgstahler and Eames 2003, among others). The final sample used for specific analyses varies due to additional data requirements and is discussed in the empirical analysis sections.

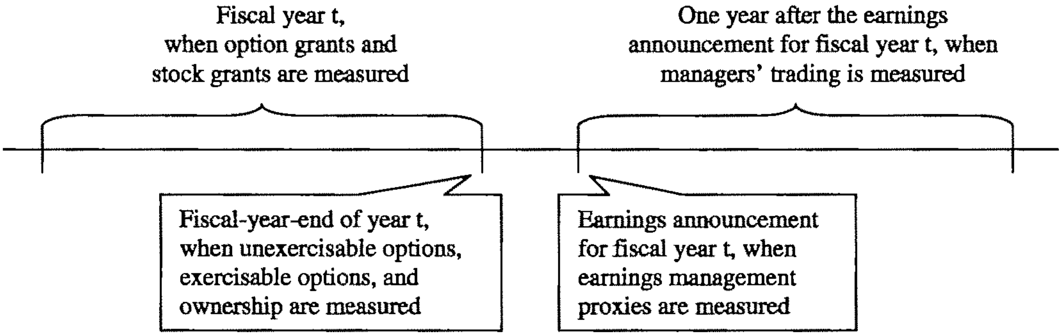
The timing of variable measurement is chosen to be consistent with our theoretical development. As summarized in Figure 3, all equity incentive variables are measured during or at the end of fiscal year *t*. Since we hypothesize that these variables are correlated with *future* manager trading and earnings management, earnings management is measured based

---

[7] Incentives to manage earnings are reduced by the associated costs, such as those arising from audit qualifications, SEC enforcement sanctions, or management reputation effects. In this paper, we first assume that the costs of earnings management are not correlated with equity incentives and thus do not affect our prediction. In Section V, we discuss the possibility that certain costs of earnings management—its impact on the value of managers' equity holdings in the future—vary with equity incentives and explicitly test the implications of this possibility.

[8] Although ExecuComp started coverage in 1992, the coverage is lower in 1992 than in other years. Thus, including 1992 data, while yielding similar results, might introduce a sample selection bias.

FIGURE 3
Timing of Variable Measurement


Fiscal year t, when option grants and stock grants are measured
One year after the earnings announcement for fiscal year t, when managers' trading is measured
Fiscal-year-end of year t, when unexercisable options, exercisable options, and ownership are measured
Earnings announcement for fiscal year t, when earnings management proxies are measured

This figure describes the timeline when the three sets of variables of interest are measured. In brief, equity incentives are measured before earnings management proxies (earnings surprises and abnormal accruals), which are based on the information disclosed at earnings announcements. Managers' trading is measured after earnings announcements.

on the information disclosed at the earnings announcement for fiscal year *t*, generally from one to four months *after* the fiscal year-end. All trading variables are measured *after* earnings announcements (subsequent to the measurement of equity incentive and earnings management variables).

Equity Incentives

We consider five equity incentive elements: option grants in the current period, unexercisable options (excluding option grants in the current period), exercisable options, restricted stock grants, and stock ownership. We deflate these measures (in shares) by total outstanding shares of the firm. Because the potential benefit of increasing short-term stock prices, if any, is shared by all shareholders (regardless of whether non-CEO shareholders realize it), the benefit enjoyed by CEOs is thus proportional to the ratio of equity incentive measures in shares to total outstanding shares.[9]

Panel A of Table 1 reports the descriptive statistics on CEO equity incentive measures. Because we assume that CEOs can sell already held shares to diversify risk due to option or stock grants, we restrict our analyses to firm-years with ownership higher than option and stock grants.[10] As reported in the table, option grants are on average 0.163 percent of outstanding shares. Option grants are part of the unexercisable options at the end of the year; the remaining unexercisable options average 0.254 percent. Exercisable options are on average 0.637 percent.

---

[9] Baker and Hall (2004) analyze the conditions under which alternative proxies for managers' equity incentives are valid. In brief, equity incentive proxies measured in proportion to firm value are most appropriate when managers' actions affect firm value on a dollar-to-dollar basis (e.g., purchasing a corporate jet). Thus, if the earnings management activities affect firm value on a dollar-to-dollar basis, instead of on a proportional basis (e.g., a corporate reorganization), the ratio measure we use is appropriate. Otherwise, a measure based on the magnitude of equity incentives is appropriate. We examine alternative measurements of equity incentives as part of our sensitivity tests reported later.

[10] This restriction increases the power of the tests but reduces the sample by about 25 percent. Relaxing this restriction does not affect the inferences.

TABLE 1
Descriptive Statistics on Equity Incentives, CEOs' Trading, and Earnings Surprises

Panel A: Descriptive Statistics on CEOs' Equity Incentives

The descriptive statistics are based on 9,472 firm-years with ownership higher than option and stock grants in the period 1993–2000. Equity incentive measures (in shares) are scaled by total outstanding shares of the firm.

	Mean	Std.	Median
Option grants (%)	0.163	0.309	0.051
Unexercisable options (excluding option grants) (%)	0.254	0.429	0.078
Exercisable options (%)	0.637	0.932	0.295
Restricted stock grants (%)[a]	0.008	0.031	0.000
Ownership (%)	4.184	7.602	0.825

Panel B: Spearman Correlations between Equity Incentive Measures

The correlations are based on 9,472 firm-years with ownership higher than option and stock grants in the period 1993–2000. Equity incentive measures (in shares) are scaled by total outstanding shares of the firm. All correlations are significant at the 0.05 level or better.

	Option Grants	Unexercisable Options (excluding option grants)	Exercisable Options
Unexercisable options (excluding option grants)	0.210		
Exercisable options	0.366	0.432	
Ownership	0.163	0.034	0.101

Panel C: CEOs' Net Sales in the Year after the Earnings Announcement

The descriptive statistics are based on all firm-years in the period 1993–2000 with both equity incentives and net sales in a specific period after the earnings announcement.

		In Million Dollars			As Percentage of Market Value		
	n	Mean	Std.	Median	Mean	Std.	Median
1st six-month	2,830	2.358	11.115	0.065	0.225	1.325	0.008
2nd six-month	2,411	2.055	9.596	0.009	0.249	1.448	0.004

Panel D: Number of Observations in Each Earnings Surprise Cell

Earnings surprise per share is measured as the difference between actual earnings and the latest consensus forecast before earnings announcement. The sample includes 4,301 firm-years with equity incentives and earnings surprises in the period 1993–2000.

Earnings Surprise (cents)	n	Earnings Surprise (cents)	n
		0	1,057
−1	379	1	746
−2	174	2	448
−3	94	3	271
−4	71	4	176
Less than −5	352	Greater than 5	533

---

[a] Due to its small magnitude, we do not analyze restricted stock grants separately from ownership in the following analyses.

Restricted stock grants appear to play a small role in stock-based compensation. While 63.2 percent of the sample report option grants, only 18.3 percent have restricted stock grants. The mean restricted stock grant is only 0.008 percent. Due to this small magnitude, stock grants can hardly have an economically significant impact on future trading or earnings management; thus for the sake of brevity, in the analyses we do not separate them from ownership, which averages 4.184 percent. These descriptive statistics are similar to those reported in Ofek and Yermack (2000).

Panel B of Table 1 reports correlations between equity incentive measures. The three option measures are positively correlated with each other and are also positively correlated with ownership.

CEOs' Trading

CEOs' trading data are collected from the SEC ownership reporting system data file (Forms 3, 4, 5). CEOs and other insiders are required to report their transactions by the tenth day of the calendar month after the trading month. As in prior research of insider trading (e.g., Lakonishok and Lee 2001), we examine three primary types of transactions: open market purchases, purchase of shares through the exercise of options, and open market sales (which include sales of shares from exercised options). Because the database assigns different names to CEOs, we regard individuals with titles like "Chairman of Board" (42.4 percent), "CEO" (12.6 percent), and "President" (45.0 percent) as CEOs, as done in Ke et al. (2003).

We examine CEOs' trading in the year after the earnings announcement for fiscal year t. CEOs' trading is measured as net sales:[11]

$$\textit{Net sales} = \textit{Open market sales} - (\textit{Open market purchases} + \textit{Options exercised}).$$

Since the frequency of CEO trading is generally low, using past trading data as the deflator, as in Lakonishok and Lee (2001), is not feasible. Instead, we scale net sales in dollars by the firm's market value at the end of fiscal year t (before net sales are measured). Using unscaled measures, either logarithm transformations or decile rankings, does not affect our inferences.

Panel C of Table 1 reports descriptive statistics on CEOs' net sales. To allow for variation in CEO trading within the year, we separately investigate net sales in the first and second six-month periods. Net sales, on average, are approximately $2 million in a six-month period, representing about 0.2 percent of the firm's market value. These descriptive statistics are consistent with those reported in Lakonishok and Lee (2001) and suggest that CEOs are net sellers on average.

Earnings Surprises

Earnings surprises are calculated as the difference between actual earnings and analysts' forecasts, both of which are measured on a per-share basis and are collected from First Call. To better capture the market's expectation, we use the most recent consensus forecasts within the three-month period before the earnings announcement. As reported in Panel D of Table 1 and consistent with prior research, there are more firm-years meeting or just beating analysts' forecasts than just missing analysts' forecasts.[12] Of 4,301 firm-years with equity incentives and earnings surprises in the period 1993–2000, 25 percent have zero earnings surprises, i.e., meeting analysts' forecasts, 17 percent beat analysts' forecasts by one cent, but less than 9 percent miss analysts' forecasts by one cent.

---

[11] Excluding *options exercised* from the calculation of net sales does not affect the inferences.

IV. PRIMARY EMPIRICAL RESULTS

As discussed in Section II, the link between equity incentives and future CEOs' trading is critical to our hypothesis that the incidence of earnings management is positively related with equity incentives. Since prior studies do not provide direct evidence on this link, in this section, we first report results on the relation between equity incentives and future trading.[13] We then report results on the impact of equity incentives on the probability of meeting or just beating analysts' forecasts. Last, we investigate whether managers sell more after earnings management.

Equity Incentives and Future Trading

We examine the relation between equity incentives (*Eq_Incent*) and future trading (*NetSale*) using the following regression:

$$NetSale_{i,t+j} = \gamma_0 + \beta Eq\_Incent_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \xi_{i,t+j}, \qquad (1)$$

where $j = 1$ or 2, denotes the jth six-month period after the earnings announcement. See Appendix A for variable measurement. Our focus is whether the coefficients on *Eq_Incent* are positive. In the regression, we also include several control variables based on prior research of insider trading.[14] Lakonishok and Lee (2001) find that insiders in large firms sell more shares, although in a smaller proportion to market value. Rozeff and Zaman (1998) find that insiders in growth firms sell more shares. Prior returns are included to control for contrarian trading behavior by insiders; insiders tend to sell more shares after high stock returns (Lakonishok and Lee 2001). Change in operating cash flow is included to control for any potential impact of firm performance on insider trading. We use a cash flow measure because it is less subject to earnings management compared to a performance measure based on earnings. The correlations reported in Appendix B indicate that these control variables are significantly correlated with each other and thus it is important to control for them at the same time.

The regression results are reported in Table 2. Consistent with our prediction, option grants, other unexercisable options, and ownership have significantly positive coefficients, suggesting that CEOs with high equity incentives are more likely to sell in the next year. Since option grants and other unexercisable options are similarly correlated with future

---

[12] Baber and Kang (2002) find that analysts' forecast data in I/B/E/S or First Call are subject to split adjustment bias; because these databases adjust stock splits retrospectively and split-adjusted data are rounded to the nearest cent, the probability of zero or one cent forecast errors might be overstated. While the impact of this split-adjustment bias on our results is unclear *ex ante*, we replicate our analyses on observations without stock splits in the sample period (61 percent of the sample). The results are essentially the same.

[13] While Ofek and Yermack (2000) study the relation between stock/option grants and contemporaneous insider sales, they neither examine future insider sales nor do they study option/stock ownership.

[14] In all analyses, we also include year dummies to control for potential year-specific effects. Throughout the paper, outliers, defined as observations with an absolute value of the R-student measure greater than 3, are excluded. None of the regressions are subject to multicollinearity based on the diagnostics in Belsley et al. (1980).

TABLE 2
Equity Incentives and Net Sales in the Future

This table reports the results from regressing CEOs' net sales (%) in the two six-month periods after earnings announcements on equity incentive measures (%) and control variables:

$$NetSale_{i,t+j} = \gamma_0 + \beta Eq\_Incent_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \xi_{i,t+j}, \quad (1)$$

where $j = 1$ or 2, denotes the jth six-month period after the earnings announcement. See Appendix A for variable measurement. Year dummies are included to control for year fixed effects. The regressions are estimated on all firm-years with data available in the sample period 1993–2000. The p-values in parentheses are based on one-sided t-tests.

	Predicted Signs	1st Six-Month	2nd Six-Month
Equity incentives			
Option grants	+	0.112	0.045
		(0.001)	(0.014)
Unexercisable options (excluding option grants)	+	0.091	0.101
		(0.001)	(0.001)
Ownership	+	0.016	0.014
		(0.001)	(0.001)
Exercisable options	+	−0.013	0.015
		(0.862)	(0.159)
Control variables			
Size	−	−0.027	−0.035
		(0.001)	(0.001)
Growth	−	−0.229	−0.191
		(0.001)	(0.001)
Stock returns	+	0.135	0.099
		(0.001)	(0.001)
Change in cash flow	+	−0.058	−0.174
		(0.712)	(0.916)
n		2,723	2,316
Adj. R^2		0.110	0.068

trading, we combine these two variables together in subsequent analyses for the sake of brevity. This modification does not affect our inferences.

In contrast, exercisable options have an insignificant impact on CEOs' trading. This result might seem unexpected, given the apparent similarity between unexercisable and exercisable options. However, unexercisable options differ from exercisable options in the hedge ratio, which captures CEOs' risk exposure due to option holdings. To illustrate, when options are first granted, the hedge ratio is between 0 and 1. If stock prices increase after option grants, then the hedge ratio of unexercisable options increases. To diversify the increased risk, the manager can sell shares already owned. By the time options are exercisable and in the money, the hedge ratio is approximately 1 and no longer increases with stock prices. Thus, there is no need for further risk diversification.[15] As a result, while unexercisable options are positively related to future insider sales, exercisable options are not.

---

[15] For example, suppose that the hedge ratio of options is 0.6 when granted (the sample average reported in Jensen and Murphy [1990]). Managers might sell shares to diversify the increase in risk (0.6 × number of options) when they are awarded options and sell more shares [(1 − 0.6) × number of options] before these options become exercisable.

One might argue that although exercisable options likely have a constant hedge ratio, they expose CEOs to the firm's idiosyncratic risk, just as ownership does. Unlike exercisable options, ownership is a more predominant equity incentive element (a much larger proportion of CEOs' wealth, as indicated in Panel A of Table 1) and is thus a better proxy for CEOs' overexposure to idiosyncratic risk. Nevertheless, we admit that there are other possible explanations for the insignificant results on exercisable options, a topic we leave for future research. Importantly for our subsequent analyses, the measures that predict future CEO trading should be correlated with earnings management, but those that do not predict future trading should not be correlated with earnings management. Our earnings management results are consistent with these predictions.

The control variable results are generally consistent with prior research. Size exhibits a negative association with net sales, suggesting that sales by CEOs of large firms are smaller in proportion to market value than those of small firms. Book-to-market also has a negative coefficient, suggesting that CEOs of growth firms sell more shares. The positive coefficient on stock returns is consistent with CEOs selling more shares after price increases. Change in cash flow, however, has an insignificant coefficient, although it becomes significantly positive when growth and stock returns are not included.

In sum, we find that CEOs with high unexercisable options or high stock ownership are more likely to sell in the future. This result suggests that equity incentives can cause managers' utility to be sensitive to short-term stock prices: if the short-term stock price is higher (lower), then the value of the shares to be sold, part of the management wealth, is higher (lower). The next section investigates whether this sensitivity induces a positive relation between equity incentives and earnings management, as hypothesized in H1.

Equity Incentives and Earnings Management

Tests of H1 are based on logit regressions of the probability of earnings surprises being in a certain range on equity incentive measures (*Eq_Incent*) and various control variables:

$$\begin{aligned}
Prob(Surprise_{i,t} \in \Phi_k) = {} & logit(\gamma_{0,k} + \beta_k Eq\_Incent_{i,t} + \gamma_{1,k} Size_{i,t} + \gamma_{2,k} Growth_{i,t-1} \\
& + \gamma_{3,k} Sales\_Growth_{i,t} + \gamma_{4,k} NOA_{i,t-1} + \gamma_{5,k} Shares_{i,t} \\
& + \gamma_{6,k} Litigation_{i,t} + \gamma_{7,k} ImplicitClaim_{i,t} \\
& + \gamma_{8,k} Num\_Estimate_{i,t} + \gamma_{9,k} CV\_AF_{i,t} + \gamma_{10,k} Down\_Rev_{i,t} \\
& + \zeta_{i,t}). \qquad (2)
\end{aligned}$$

$Surprise_{i,t}$ is earnings surprise per share, as defined above; Φ could be one of the following two sets of earnings surprises: (1) negative earnings surprises; (2) zero or one cent; equity incentive measures include unexercisable options, exercisable options, and ownership. See Appendix A for variable measurement. Hypothesis 1 predicts positive coefficients on equity incentive measures when investigating the probability of meeting or just beating analysts' forecasts (i.e., earnings surprises being zero or one cent). Relatedly, H1 implies a negative relation between equity incentives and the probability of missing analysts' forecasts (i.e., negative earnings surprises).

Following prior research, we include various control variables that might be correlated with earnings surprises and/or equity incentives. Equity incentives are generally higher for small firms or growth firms, and high-growth firms might be more likely to meet or beat analysts' forecasts to avoid the torpedo effect associated with missing analysts' forecasts, as documented in Skinner and Sloan (2002). Thus, we control for firm size and growth (proxied by the book-to-market ratio and sales growth). Beginning-of-period net operating assets are included because Barton and Simko (2002) find that firms with high net operating assets in the previous year are less likely to meet or just beat analysts' forecasts, arguably due to a lack of flexibility in managing earnings upward. The number of outstanding shares is included because firms with more outstanding shares have smaller earnings per share and are thus more likely to meet or just beat analysts' forecasts of earnings per share (Barton and Simko 2002).

Prior research suggests other incentives to avoid negative earnings surprises. Ali and Kallapur (2001) and Francis et al. (1994) find that price declines due to negative earnings surprises can lead to litigation liability. Similarly, Bowen et al. (1995) argue that negative earnings surprises can have negative publicity effects, which can adversely affect the implicit claims between a firm and its stakeholders. Thus, the incentive to avoid negative earnings surprises increases with litigation risk and implicit claims. We include an industry-based litigation dummy variable and an implicit claim variable (proxied by labor intensity) as controls.

We also include three variables to control for analyst forecast attributes: number of analysts, dispersion of individual forecasts, and a downward revision dummy. The incentives to meet analysts' forecasts can be stronger if more analysts are following the firm or if there is greater consensus among analysts (Payne and Robb 2000). As shown in Matsumoto (2002), it is easier to meet or beat analysts' forecasts that have been guided downward, so we include a dummy variable to indicate whether analysts' forecasts are revised downward in the three months before earnings announcements.

As reported in Appendix B, the univariate correlations between the probability of meeting or just beating analysts' forecasts and independent variables, especially the equity incentive measures, are generally consistent with our predictions. While many correlations between independent variables are significant, none is substantial enough to lead to multicollinearity.

Table 3 reports the logit regression results. As expected, both unexercisable options and ownership exhibit significant positive effects on the probability of meeting or just beating analysts' forecasts. For example, a one standard deviation increase in unexercisable options increases by 16.3 percent the odds of meeting or just beating analysts' forecasts, while a one standard deviation increase in ownership increases by 30.5 percent the odds of meeting or just beating analysts' forecasts.[16] Also as expected, unexercisable options and ownership are negatively associated with the probability of missing analysts' forecasts.

Although exercisable options are negatively correlated with the probability of missing analysts' forecasts, they are not correlated with the probability of meeting or just beating analysts' forecasts. This result is consistent with the trading evidence reported in the last section and with the argument that equity incentives lead to earnings management via future trading. Unlike unexercisable options and ownership, exercisable options are not correlated with future trading and are thus not correlated with earnings management in this test or our other analyses.

---

[16] The odds of an event are the ratio of the probability that this event occurs to the probability that this event does not occur. For example, the odds of meeting or just beating analysts' forecasts are the ratio of the probability that a firm meets or just beats analysts' forecasts to the probability of the firm having earnings surprises other than zero or one cent. The odds are a positive function of the probability that the event occurs. We interpret the results based on the change in the odds because this measure is a function of the coefficient in the logit regressions only. In contrast, calculating the change in the probability requires a reference point.

TABLE 3
Equity Incentives and Meeting or Just Beating Analysts' Forecasts

The table reports the change in the odds of firms (1) missing analysts' forecasts, or (2) meeting or just beating analysts' forecasts by one cent, due to a change of one standard deviation of a continuous variable or from 0 to 1 for a dummy variable, and the p-value (in parentheses) of the corresponding coefficient in the logit regressions.

$$\begin{aligned} Prob(Surprise_{i,t} \in \Phi_k) = logit(&\gamma_{0,k} + \beta_k Eq\_Incent_{i,t} + \gamma_{1,k} Size_{i,t} + \gamma_{2,k} Growth_{i,t-1} \\ &+ \gamma_{3,k} Sales\_Growth_{i,t} + \gamma_{4,k} NOA_{i,t-1} + \gamma_{5,k} Shares_{i,t} \\ &+ \gamma_{6,k} Litigation_{i,t} + \gamma_{7,k} ImplicitClaim_{i,t} + \gamma_{8,k} Num\_Estimate_{i,t} \\ &+ \gamma_{9,k} CV\_AF_{i,t} + \gamma_{10,k} Down\_Rev_{i,t} + \zeta_{i,t}) \end{aligned} \tag{2}$$

$Surprise_{i,t}$ is earnings surprises per share; Φ could be one of the following two sets of earnings surprises: (1) negative earnings surprises; (2) zero or one cent. See Appendix A for variable measurement. Year dummies are included to control for year fixed effects. The logit regressions are based on 4,301 firm-years with all required data in the period 1993–2000.

	Negative Earnings Surprises $\phi = (-\infty, 0)$ (1)		**Meeting or Just Beating Analysts' Forecasts $\phi = [\$0, \$0.01]$ (2)**	
	Predicted Signs	**Change in the Odds (p-value)**	**Predicted Signs**	**Change in the Odds (p-value)**
Equity incentives				
Unexercisable options	–	–0.159	+	0.163
		(0.001)		(0.001)
Ownership	–	–0.224	+	0.305
		(0.001)		(0.001)
Exercisable options	–	–0.098	+	–0.043
		(0.010)		(0.869)
Control variables				
Size	?	–0.352	?	0.032
		(0.001)		(0.607)
Growth	+	–0.090	–	–0.243
		(0.985)		(0.001)
Sales_Growth	–	–0.194	+	–0.033
		(0.001)		(0.824)
Net operating assets (NOA)	+	0.100	–	0.025
		(0.010)		(0.737)
Shares	–	0.016	+	0.309
		(0.599)		(0.001)
Litigation	–	–0.083	+	0.589
		(0.160)		(0.001)
Implicit Claim	–	–0.063	+	0.187
		(0.052)		(0.001)
Num_Estimate	–	–0.121	+	0.242
		(0.012)		(0.001)
CV_AF	+	0.496	–	–0.733
		(0.001)		(0.001)
Down_Rev	–	7.939	+	0.980
		(0.999)		(0.406)
Pseudo R^2		0.195		0.180
Likelihood ratio		603		612
(p-value)		(0.001)		(0.001)

Results for the control variables are generally consistent with those reported in prior research. Large firms are less likely to miss analysts' forecasts and growth firms are more likely to meet or just beat analysts' forecasts. Consistent with Barton and Simko (2002), firms with large net operating assets in the previous year are more likely to miss analysts' forecasts. Firms with more outstanding shares, firms in high-litigation industries, and firms with more implicit claims are more likely to meet or just beat analysts' forecasts. Also, firms are more likely to meet or just beat analysts' forecasts when more analysts provide forecasts or when there is greater consensus among forecasts. In contrast to the expectation management argument, firms with downward revisions are more likely to miss analysts' forecasts than those without downward revisions, but this result is consistent with analysts' under-reaction to bad news, as documented in Abarbanell and Bernard (1992), because the downward revision can be due to underlying economic changes.

To further investigate whether expectation management contributes to the reported results, we replicate the analyses using "old" consensus forecasts—the latest consensus forecast six months before the earnings announcement. These "old" forecasts are less subject to management guidance, which generally occurs in a short period before earnings announcements. If expectation management is driving the results, then we should not observe similar patterns when we use "old" consensus forecasts. The results, not reported for the sake of brevity, are similar to those reported, suggesting that our results are mainly due to earnings management.[17]

To summarize, the results presented in this section are consistent with H1. CEOs with high equity incentives are more likely to meet or just beat analysts' forecasts and less likely to miss analysts' forecasts, compared to those with low equity incentives.

Earnings Management and Future Trading Conditional on Equity Incentives

In this section, we investigate the impact of meeting or beating analysts' forecasts on net sales separately for managers with high equity incentives and for those with low equity incentives. Observing more selling for managers with high equity incentives, who are more likely to manage earnings based on results reported above, suggests that managers increase sales after earnings management.

To capture the differential selling, we first classify the sample into four groups based on aggregate equity incentives and earnings surprises:

H_MB = firms with ***high*** equity incentives and ***meeting or beating*** analysts' forecasts;
H_Miss = firms with ***high*** equity incentives and ***missing*** analysts' forecasts;
L_MB = firms with *low* equity incentives and ***meeting or beating*** analysts' forecasts; and
L_Miss = firms with *low* equity incentives and ***missing*** analysts' forecasts.

We measure the aggregate equity incentives as the sum of unexercisable options, weighted by their hedge ratios as measured in Core and Guay (1999), and ownership; high (low) equity incentive managers are those with aggregate equity incentives higher (lower) than the sample median.

We then include dummy variables that indicate each of the four groups into Equation (1):

---

[17] The overall weak results on expectation management are consistent with its ambiguous net effect on stock prices. Managers can benefit from the guidance if the stock price is higher after (1) managing analysts' forecasts downward (and presumably resulting in a lower price) and (2) meeting or beating analysts' forecasts later. If managers are uncertain whether the upward impact of meeting or beating analysts' forecasts outweighs the downward impact of guidance, then they are likely not to use this tool.

$$NetSale_{i,t+j} = \beta_1 H\_MB_{i,t} + \beta_2 H\_Miss_{i,t} + \beta_3 L\_MB_{i,t} + \beta_4 L\_Miss_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \xi_{i,t+j}. \quad (3)$$

The coefficient on each dummy variable captures how much each group of firms sells in the future. The difference in coefficients, $\beta_1 - \beta_2$ ($\beta_3 - \beta_4$), indicates whether high (low) equity incentive managers sell more after meeting or beating analysts' forecasts than after missing analysts' forecasts.

Table 4 reports the regression results. Coefficients on all dummy variables are significantly positive, suggesting that firms in each group are net sellers. As reported at the bottom of the table, the coefficient on *H_MB* is significantly higher than that on *H_Miss*, at the 0.05 level in the first six-month period and at the 0.08 level in the second six-month period. Thus, for high equity incentive managers, meeting or beating analysts' forecasts leads to more net sales in the first year after earnings announcements than missing analysts' forecasts. However, meeting or beating analysts' forecasts does not appear to motivate low equity incentive managers to sell more: there is no significant difference between the coefficient on *L_MB* and that on *L_Miss*.[18]

Overall, these results are consistent with high equity incentive managers increasing their sales after earnings management. Doing so is beneficial if such earnings management can increase stock prices in the near term. In this sense, subsequent trading serves as the "exit strategy" for earnings management, thereby increasing the credibility of our earnings management evidence (Schipper 1989).

V. ADDITIONAL ANALYSES

Equity Incentives and Earnings Smoothing

The evidence that managers with high equity incentives are more likely to report earnings that meet or just beat analysts' forecasts is consistent with their incentives to manage earnings upward to avoid negative earnings surprises, as suggested by Stein (1989). At the same time, managers might have incentives to manage earnings downward in years with good performance to reserve for the future. Earnings smoothing is appealing in this context for at least two important reasons. First, it is difficult, if not impossible, for a firm to manage earnings upward consistently. By engaging in income-decreasing earnings management in years with good performance, managers can increase earnings as needed in the future. Such smoothing reduces the likelihood of missing analysts' forecasts and increases the likelihood of meeting analysts' forecasts in future periods.

Second, equity incentives can be recurring—managers with high equity incentives are more likely to have high equity incentives in the future. Unreported results indicate that equity incentives are positively serially correlated. Relatedly, as reported in Panel A of Table 5, equity incentives are positively correlated with managers' sales in the second year after earnings announcements. (As before, exercisable options are not correlated with future trading.) Thus, managers with high equity incentives also care about stock prices in the future; if they cannot meet analysts' forecasts in the future, then they bear negative consequences.

These discussions suggest that high equity incentive managers are likely to smooth earnings. One way to smooth earnings is to avoid large positive earnings surprises. By doing so, managers can reserve for the future and reduce the likelihood of missing analysts' forecasts in the future. Thus, we expect a negative association between the incidence of

---

[18] A formal F-test indicates that ($\beta_1 - \beta_2$) − ($\beta_3 - \beta_4$) is significantly different from zero.

TABLE 4
The Impact of Meeting or Beating Analysts' Forecasts on CEOs' Net Sales Conditional on Equity Incentives

This table reports results from regressing future net sales (%) on four dummy variables, *H_MB*, *H_Miss*, *L_MB*, and *L_Miss*, and control variables:

$$NetSale_{i,t+j} = \beta_1 H\_MB_{i,t} + \beta_2 H\_Miss_{i,t} + \beta_3 L\_MB_{i,t} + \beta_4 L\_Miss_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \zeta_{i,t+j}, \quad (3)$$

where $j = 1$ or 2, denotes the *j*th six-month period after the earnings announcement, and the four dummy variables are defined as follows:

H_MB = 1 if the firm has high equity incentives and meets or beats analysts' forecasts;
H_Miss = 1 if the firm has high equity incentives and misses analysts' forecasts;
L_MB = 1 if the firm has low equity incentives and meets or beats analysts' forecasts; and
L_Miss = 1 if the firm has low equity incentives and misses analysts' forecasts.

High or low equity incentives are classified based on the sample median of an aggregate equity incentive measure, which is the sum of unexercisable options, weighted by their hedge ratios as measured in Core and Guay (1999), and ownership. See Appendix A for the measurement of other variables. Year dummies are included to control for year fixed effects. The regressions are estimated on all firm-years with data available in the period 1993–2000. The p-values in parentheses are based on one-sided t-tests. The table also reports the difference in coefficients between *H_MB* and *H_Miss*, that between *L_MB* and *L_Miss*, and the p-value of the corresponding F-tests.

	Predicted Signs	1st Six-Month	2nd Six-Month
H_MB (β_1)	+	0.620 (0.001)	0.567 (0.001)
H_Miss (β_2)	+	0.534 (0.001)	0.500 (0.001)
L_MB (β_3)	+	0.469 (0.001)	0.414 (0.001)
L_Miss (β_4)	+	0.419 (0.001)	0.410 (0.001)
Control variables		Yes	Yes
n		1,637	1,389
Adj. R^2		0.168	0.158
The impact of meeting or beating analysts' forecasts on net sales (p-value)			
For high equity incentive managers ($\beta_1 - \beta_2$)	+	0.086 (0.054)	0.067 (0.077)
For low equity incentive managers ($\beta_3 - \beta_4$)	+	0.050 (0.186)	0.004 (0.908)

large positive earnings surprises and equity incentives, and this negative association should be stronger for firms with persistent equity incentives, which arguably have stronger incentives to smooth earnings. We test these two predictions below.

Equity Incentives and Avoiding Large Positive Earnings Surprises

Column (1) of Panel B, Table 5 reports the logit regression results of the probability of large positive earnings surprises (i.e., earnings surprises equal to or greater than four cents) on equity incentives and control variables as in Equation (2). The results are consistent with earnings smoothing. A one standard deviation increase in unexercisable options reduces by 14.3 percent the odds of having large positive earnings surprises, while a one standard deviation increase in ownership reduces by 11.6 percent the odds of having large positive earnings surprises. The results are similar if we use different cutoff points (such as three cents, five cents, or six cents) to measure large positive earnings surprises. Consistent with the future trading evidence, exercisable options are not negatively correlated with the probability of large positive earnings surprises.

If reserving for the future is successful, equity incentives should be positively correlated with the probability of meeting or just beating analysts' forecasts in the next year. Column (2) of Panel B reports the logit regression of meeting or just beating analysts' forecasts in year $t+1$ on equity incentives in year t. While the results for option holdings are insignificant, firms with high ownership are more likely to meet or just beat analysts' forecasts in the next year, consistent with earnings smoothing.

TABLE 5
Equity Incentives and Earnings Smoothing

Panel A: Equity Incentives and CEOs' Trading in the Second Year after Earnings Announcements

This panel reports the results from regressing CEOs' net sales (%) in the second year (the 3rd and 4th six-month periods) after earnings announcements on equity incentive measures (%) and control variables:

$$NetSale_{i,t+j} = \gamma_0 + \beta\ Eq\_Incent_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \zeta_{i,t+j}, \quad (1)$$

where $j = 3$ or 4, denotes the jth six-month period after the earnings announcement. See Appendix A for variable measurement. The regressions are estimated on all firm-years with data available in the period 1993–2000. Year dummies are included to control for year fixed effects. The p-values in parentheses are based on one-sided t-tests.

	Predicted Signs	3rd Six-Month	4th Six-Month
Equity incentives			
Unexercisable options	+	0.076	0.068
		(0.002)	(0.011)
Ownership	+	0.021	0.019
		(0.001)	(0.001)
Exercisable options	+	0.030	−0.002
		(0.144)	(0.538)
Control variables			
Size	−	−0.037	−0.048
		(0.001)	(0.001)
Growth	−	−0.233	−0.256
		(0.001)	(0.001)
Stock returns	+	0.052	0.079
		(0.026)	(0.007)
Change in cash flow	+	0.128	−0.191
		(0.790)	(0.867)
n		2,213	1,624
Adj. R^2		0.085	0.086

(*continued on next page*)

TABLE 5 (continued)

Panel B: Equity Incentives and Earnings Smoothing

This panel reports the change in the odds of firms (1) beating analysts' forecasts by at least four cents in the current year, and (2) meeting or just beating analysts' forecasts by one cent in the next year, due to a change of one standard deviation of a continuous variable, and the p-value (in parentheses) of the corresponding coefficient in the logit regressions. See Table 3 for model specifications and Appendix A for variable measurement. Year dummies are included to control for year fixed effects. Column (1) is based on current year's earnings surprises ($Surprise_{i,t}$) and Column (2) is based on the next year's earnings surprises ($Surprise_{i,t+1}$). $Surprise_{i,t+1}$ is the difference between First Call actual EPS for year $t+1$ and the latest consensus forecast for year $t+1$ disclosed before the earnings announcement for year $t+1$, rounded to the nearest cent.

	The Probability of Large Positive Earnings Surprises in the *Current* Year Prob($Surprise_{i,t} \geq \$0.04$) (1)		**The Probability of Meeting or Just Beating Analysts' Forecasts in the *Next* Year Prob($Surprise_{i,t+1} = \$0$ or $\$0.01$) (2)**	
	Predicted Signs	**Change in the Odds (p-value)**	**Predicted Signs**	**Change in the Odds (p-value)**
Equity incentives				
Unexercisable options	–	−0.143	+	0.041
		(0.001)		(0.170)
Ownership	–	−0.116	+	0.167
		(0.007)		(0.001)
Exercisable options	–	0.152	+	−0.064
		(0.998)		(0.942)
Control variables		Yes		Yes
n		4,301		3,429
Pseudo R^2		0.150		0.135
Likelihood ratio		397		291
(p-value)		(0.001)		(0.001)

Panel C: The Incremental Impact of Persistent Equity Incentives on Earnings Smoothing

This panel reports the incremental impact of persistent equity incentives on earnings surprises: the probability of large positive earnings surprises in Column (1) and the probability of meeting or just beating analysts' forecasts in Column (2), using the following regression:

$$Prob(Surprise_{i,t} \in \Phi_k) = logit(\gamma_{0,k} + \beta_{1,k}Total_{i,t} + \beta_{2,k}Total_{i,t} \times Recurring_i + \beta_{3,k}Exercisable\ options_{i,t} + \gamma_k Controls_{i,t} + \zeta_{i,t}),$$

where *Controls* are those included in Equation (2). The table reports the change in the odds due to a change of one standard deviation of a variable and the p-value (in parentheses) of the corresponding coefficient in the logit regressions. The aggregate equity incentive measure, *Total*, is the sum of unexercisable options, weighted by the hedge ratio as measured in Core and Guay (1999), and ownership. To identify firms with persistent equity incentives, we first rank all firm-years into three equal-sized groups: high, medium, or low, based on the aggregate equity incentive measure. We then calculate the average ranking for each firm over the sample period. If a firm's average ranking is above the 3rd quartile or below the 1st quartile of the average ranking distribution, this firm is regarded as having persistent equity incentives (consistently high or consistently low equity incentives, respectively); for these firms, the dummy variable, *Recurring*, is defined as 1, and 0 otherwise. See Appendix A for the measurement of other variables. Year dummies are included to control for year fixed effects. The logit regressions are based on 3,958 firm-years with all required data in the period 1993–2002.

(*continued on next page*)

TABLE 5 (continued)

	Large Positive Earnings Surprises $\phi = [\$0.04, \infty)$ (1)		Meeting or Just Beating Analysts' Forecasts $\phi = [\$0, \$0.01]$ (2)	
	Predicted Signs	**Change in the Odds (p-value)**	**Predicted Signs**	**Change in the Odds (p-value)**
Total	?	0.024 (0.760)	+	0.262 (0.001)
Total × *Recurring*	–	−0.218 (0.001)	?	0.038 (0.509)
Exercisable options	?	0.063 (0.233)	?	0.002 (0.961)
Control variables		Yes		Yes
Pseudo R^2		0.164		0.198
Likelihood ratio (p-value)		368 (0.001)		577 (0.001)

Earnings Smoothing for Firms with Persistent Equity Incentives

To test whether firms with persistent equity incentives are more likely to smooth earnings, we first rank all firm-years into three equal-sized groups: high, medium, or low, based on the aggregate equity incentive measure—the sum of unexercisable options, weighted by the hedge ratio as measured in Core and Guay (1999), and ownership. We then calculate the average ranking for each firm over the sample period. To increase the reliability of the classification, we only consider the 1,400 firms that have at least five years of data from ExecuComp over the eight-year period 1993–2000.

If a firm's average ranking is above the third quartile or below the first quartile of the average ranking distribution, then this firm is regarded as having persistent equity incentives (consistently high or consistently low equity incentives, respectively). Under this approach, 56 percent of the firm-years exhibit persistent equity incentives. We use a dummy variable, *Recurring*, to denote these firm-years.[19] We then add an interaction of this dummy with equity incentives to the logit regression (Equation (2)) to capture the incremental effect for firms with persistent equity incentives. To simplify the analyses, we use the aggregate equity incentive measure (*Total*).

The results are reported in Column (1) of Panel C, Table 5. As expected, equity incentives are negatively correlated with the probability of large positive earnings surprises only for firms with persistent equity incentives. That is, firms with persistent equity incentive measures are more likely to reserve for the future and those with less persistent equity incentives do not have such incentives.

In addition, we test whether managers with persistent equity incentives drive the results on meeting or just beating analysts' forecasts reported in Table 3. In Column (2) of Panel C, we add an interaction between the recurring dummy and equity incentives in the logit regression of meeting or just beating analysts' forecasts in the current period. The coefficient on the interaction term is insignificant, suggesting that equity incentives affect the likelihood of meeting or just beating analysts' forecasts similarly for firms with persistent or non-persistent equity incentives.

---

[19] The results are similar if we exclude firm-years with consistently low equity incentives.

Equity Incentives and Earnings Management via Abnormal Accruals

In this section, we investigate the relation between equity incentives and abnormal accruals, a commonly used proxy for earnings management. In this paper's context, managers can benefit from income-increasing abnormal accruals. At the same time, the earnings smoothing incentives documented above imply income-decreasing abnormal accruals in years with good performance. Thus, *ex ante*, the direction of the correlation between equity incentives and abnormal accruals is unclear. If the incentive for increasing income dominates that for reserving for the future, the correlation is positive; otherwise, it is negative.

Abnormal accruals are defined as the difference between total accruals and normal accruals. Normal accruals are estimated using the cross-sectional Jones (1991) model as described in DeFond and Jiambalvo (1994). Panel A of Table 6 describes the estimation method and reports the descriptive statistics on abnormal accruals. The mean and median abnormal accruals (scaled by beginning-of-period total assets) are 0.002 and 0.001, respectively. Untabulated statistics suggest that these numbers are insignificantly different from zero at the 0.05 level.

Panel A of Table 6 also compares abnormal accruals and other firm characteristics between our sample and all Compustat firms with abnormal accruals. While there is no significant difference in abnormal accruals, the sample firms are larger and have better performance than the Compustat population. The average total assets are around $2 billion for the sample firms, but only $886 million for the Compustat population. Earnings and operating cash flow on average are 5 percent and 11.5 percent of beginning-of-period assets for the sample firms, but −10.8 percent and −2.6 percent for the Compustat population, respectively. These differences are due to data restrictions on equity incentive measures and we control for both size and firm performance in the analyses.[20]

Equity Incentives and Abnormal Accruals

We use the following regression to investigate the impact of equity incentives (*Eq_Incent*) on abnormal accruals:

$$\begin{aligned} Abnormal\ Accruals_{i,t} = {} & \gamma_0 + \beta\ Eq\_Incent_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Leverage_{i,t} \\ & + \gamma_3 Bonus_{i,t} + \gamma_4 Risk_{i,t} + \gamma_5 Growth_{i,t} + \gamma_6 NOA_{i,t-1} \\ & + \gamma_7 Litigation_{i,t} + \gamma_8 ImplicitClaim_{i,t} + \zeta_{i,t} \qquad (4) \end{aligned}$$

See Appendix A for variable measurement. In the regression, we control for factors that might affect abnormal accruals based on prior research. Positive accounting theory suggests that managers tend to manage earnings to decrease political costs (proxied by size), to relax debt covenants (the closeness to debt covenant violation is proxied by leverage), and to increase bonuses (Healy 1985; Watts and Zimmerman 1986; Sweeney 1994; among others).[21] In addition, firms with high equity incentives are generally faster-growing and riskier (Lambert and Larcker 1987; Murphy 1999), and Warfield et al. (1995) indicate that riskier and high-growth firms have more abnormal accruals (which could be of either sign). Thus, we also control for the impact of systematic risk and growth opportunities, and several other variables that might affect earnings management, as discussed in Section IV.

---

[20] ExecuComp covers companies within the S&P 1500 (the combination of the S&P 500 index, the S&P MidCap 400 index, and the S&P SmallCap 600 index). In addition to size requirements, companies must have positive earnings (excluding discontinued operations and extraordinary items) to be included in these indexes.

[21] To reduce the burden of data collection, details of bonus plans are not exploited in this study. Instead, we use a dummy variable to indicate firm-years with nonzero bonus. We find similar results when using alternative measures of bonus, including a dummy variable based on the ranking of return on assets as used in McNichols and Wilson (1988) and Bowen et al. (1995) and the ratio of cash bonus over total compensation. We also estimate the regressions after excluding the bonus variable and find similar results on equity incentive variables.

The regression results are reported in Panel B of Table 6. Both unexercisable options and ownership are positively correlated with abnormal accruals, and exercisable options are not. These results suggest that *on average*, high equity incentive CEOs increase earnings

TABLE 6
Equity Incentives and Earnings Management via Abnormal Accruals

Panel A: Descriptive Statistics on Abnormal Accruals and Other Firm Characteristics

	The Sample (n = 6,867)			Compustat Population (n = 50,984)		
	Mean	**Median**	**Inter-Quartile Range**	**Mean**	**Median**	**Inter-Quartile Range**
Abnormal accruals	0.002	0.001	0.086	0.007	0.006	0.141
Total assets ($million)	2,028	624	1,537	886	66	310
Earnings	0.050	0.068	0.087	−0.108	0.026	0.196
Operating cash flows	0.115	0.113	0.111	−0.026	0.056	0.191

All variables, except total assets, are deflated by beginning-of-period total assets. The descriptive statistics for the sample are based on all firm-years with both equity incentives and abnormal accruals in the period 1993–2000. The Compustat population includes all firm-years in the same period with abnormal accruals. Abnormal accruals are defined as the difference between total accruals and normal accruals. Normal accruals are estimated using the cross-sectional Jones (1991) model as described in DeFond and Jiambalvo (1994). In this model, normal accruals are estimated as a function of change in revenues and level of property, plant, and equipment:

$$TA_{i,t}/A_{i,t-1} = b_0(1/A_{i,t-1}) + b_1(\Delta REV_{i,t}/A_{i,t-1}) + b_2(PPE_{i,t}/A_{i,t-1}) + \mu_{i,t}$$

where:

$TA_{i,t}$ = total accruals, i.e., income before extraordinary items (#18) minus operating cash flows (#308);
$A_{i,t-1}$ = beginning-of-period total assets (#6);
$\Delta REV_{i,t}$ = change in net revenues (#12) from year $t-1$ to year t, included to control for normal changes in working capital;
$PPE_{i,t}$ = gross property, plant, and equipment (#7), included to control for normal depreciation expense;
$\mu_{i,t}$ = error term;
i,t = subscript for firm i in year t.

The portion of total accruals that is not explained by normal operating activities, i.e., the residual $\mu_{i,t}$, is the estimate of abnormal accruals. Because operating characteristics might vary across industries and years, the estimation is conducted for each industry-year combination, based on all firm-years with data available from Compustat. Industry is defined on the basis of two-digit SICs and industry-year combinations with less than seven observations are deleted.

(*continued on next page*)

TABLE 6 (continued)

Panel B: Regressions of Abnormal Accruals on Equity Incentive Measures and Control Variables

This panel reports the coefficients and the accompanying p-value of the following regression:

$$\begin{aligned} Abnormal\ Accruals_{i,t} = {} & \gamma_0 + \beta\ Eq\_Incent_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Leverage_{i,t} + \gamma_3 Bonus_{i,t} \\ & + \gamma_4 Risk_{i,t} + \gamma_5 Growth_{i,t} + \gamma_6 NOA_{i,t-1} + \gamma_7 Litigation_{i,t} \\ & + \gamma_8 Implicit\ Claim_{i,t} + \zeta_{i,t} \qquad (4) \end{aligned}$$

This regression is estimated using the pooled sample with year dummies to control for year fixed effects. See Appendix A for variable measurement. The p-values in parentheses are based on two-sided t-tests.

	Abnormal Accruals
Equity incentives	
Unexercisable options	0.695 (0.012)
Ownership	0.070 (0.001)
Exercisable options	0.140 (0.407)
Control variables	
Size	0.001 (0.529)
Leverage	−0.032 (0.003)
Bonus	0.007 (0.059)
Risk	−0.001 (0.841)
Growth	−0.005 (0.342)
Net operating assets	−0.014 (0.001)
Litigation	−0.013 (0.001)
Implicit claim	−0.019 (0.001)
n	6,387
Adj. R^2	0.022

(*continued on next page*)

TABLE 6 (continued)

Panel C: Regressions of CEOs' Net Sales on Abnormal Accruals

This panel reports the results from regressing future net sales (%) on equity incentives (%), abnormal accruals, and control variables:

$$NetSale_{i,t+j} = \gamma_0 + \beta\ Eq\_Incent_{i,t} + \delta\ AbnormalAccruals_{i,t} + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \xi_{i,t+j},$$

where $j = 1$ or 2, denotes the jth six-month period after the earnings announcement. Positive and negative abnormal accruals are included separately to capture their potentially differential effects. *Positive abnormal accruals* are zero when abnormal accruals are negative, and *negative abnormal accruals* are zero when abnormal accruals are positive. To reduce the impact of measurement errors in abnormal accruals, decile ranks (scaled to the range between 0 and 1) are used. See Appendix A for variable measurement. The regressions are estimated on all firm-years with data available in the period 1993–2000. Year dummies are included to control for year fixed effects. The p-values in parentheses are based on one-sided t-tests for coefficients with signed predictions and on two-sided t-tests otherwise.

	Predicted Signs	1st Six-Month	2nd Six-Month
Equity incentives			
Unexercisable options	+	0.086 (0.001)	0.064 (0.002)
Ownership	+	0.016 (0.001)	0.014 (0.001)
Exercisable options	+	−0.008 (0.749)	0.018 (0.117)
Abnormal accruals			
Positive abnormal accruals	+	0.075 (0.019)	0.074 (0.054)
Negative abnormal accruals	?	−0.050 (0.171)	−0.073 (0.093)
Control variables		Yes	Yes
n		2,723	2,316
Adj. R^2		0.109	0.068

via abnormal accruals.[22] With respect to control variables, firms with high leverage, without a bonus plan, with high beginning-of-period net operating assets, in high litigation industries, and with high implicit claims take less positive abnormal accruals.

We conduct additional analyses to assess the robustness of our results to possible measurement errors in abnormal accruals (Dechow et al. 1995; Bernard and Skinner 1996). First, because firms with extreme performance are more likely to have measurement errors, we add a control for firm performance (return on assets, i.e., *ROA*) or exclude firms with extreme earnings (top 5 percent, 10 percent, or 20 percent based on $|ROA|$ ranking). Second, we include risk, earnings persistence, and earnings variability as explanatory variables to better control for the volatility of the firm's economic environment or operations. Third, we replace abnormal accruals with inventory changes, which are shown to be used for earnings management purposes and are less subject to measurement errors than abnormal accruals (Thomas and Zhang 2002). Our results are robust to these alternative specifications.

---

[22] Warfield et al. (1995) find a negative association between managerial ownership and the absolute value of abnormal accruals. Unlike Warfield et al. (1995), we focus on the relation between all equity incentive elements and signed abnormal accruals. Thus, one should be careful in comparing our results with those of Warfield et al. (1995).

Overall, these results suggest that CEOs with high equity incentives take more income-increasing abnormal accruals than those with low equity incentives. That is, the incentive to increase earnings dominates that to reserve for the future in this particular sample. While this finding appears at odds with the earnings smoothing argument discussed above, additional analyses indicate that the results are largely driven by managers with less persistent equity incentives, who are less concerned with accrual reversal and have fewer incentives to reserve for the future. When we split the sample based on the average ranking of equity incentives over the sample period, as discussed above, we find that equity incentives are not significantly correlated with abnormal accruals for firms with persistent equity incentives, but the correlation is significantly positive for firms with less persistent equity incentives.

Abnormal Accruals and Future CEO Trading

Similar to analyzing the relation between meeting or beating analysts' forecasts and future CEO trading, we also investigate the relation between abnormal accruals and future trading. Panel C of Table 6 reports regression results when abnormal accruals are added to Equation (1). We add positive and negative abnormal accruals separately in the regression to detect any differential impact. The coefficients on positive abnormal accruals are significantly positive, and the coefficients on negative abnormal accruals are insignificant (marginally significantly negative in the second six-month period). These results are consistent with CEOs increasing sales after income-increasing abnormal accruals.[23]

Other Sensitivity Tests

Alternative Measures of Equity Incentives

We examine the robustness of our results to the measurement of equity incentives in several ways. First, in the main analysis, we implicitly assume that the hedge ratio of options is constant across firms. To relax this assumption, we explicitly measure the hedge ratio for individual firms, as done in Core and Guay (1999), and adjust the measurement of exercisable and unexercisable options accordingly (i.e., multiplying the option measures by the associated hedge ratio). Second, as argued in Baker and Hall (2004), if earnings management affects firm value on a proportionate basis rather than on a dollar-for-dollar basis, a more appropriate measure is the dollar value of the incentive variable. Thus, we also use the magnitude of equity incentives, measured as the natural logarithm of one plus the change in the value of unexercisable options and ownership due to 1 percent increase in stock prices. The results based on these alternative equity incentive measures are similar to those reported.

Last, we measure stock-based compensation as the ratio of stock grants and option grants (both in dollars) to the total compensation (the sum of salary, bonus, stock grants, option grants, long-term incentive payouts, and other compensation). This measure captures the proportion of CEOs' compensation that is sensitive to stock prices. The results on this alternative measure, as reported in Panel A of Table 7, are consistent with earnings management behavior: stock-based compensation is positively correlated with the probability of meeting or just beating analysts' forecasts. The results on other variables, including ownership, are similar to those from the main analyses.

---

[23] Further analyses (not reported for the sake of brevity) indicate that managers taking more negative abnormal accruals sell more in the second year after earnings announcements, presumably after the reversal of abnormal accruals. This result is consistent with managers postponing sales if they take negative abnormal accruals.

Analyses of Future CEO Trading Conditional on Future Realized Returns

As discussed in Section II, one important reason for managers with high equity incentives to sell shares in the future is that their exposure to the idiosyncratic risk increases with stock prices, because the sensitivity of options to stock price increases, or because their wealth becomes more concentrated in their companies' stocks, or both. If this is true, then it follows that managers with the same equity incentives will sell more when stock prices increase more. To explicitly test this prediction, we add interactions between equity incentive measures and future realized returns into the future trading regressions. We use a dummy variable to capture large increases in stock prices; the dummy variable is 1 if the realized return in the period when net sales are measured is in the top quartile.

We report the results in Panel B of Table 7. Consistent with our prediction, the interactions of unexercisable options and ownership with the high return dummy have significantly positive coefficients, suggesting that managers sell more in the future due to equity incentives when stock prices increase dramatically than otherwise. These results lend further support to our argument on future trading motivated by equity incentives.

TABLE 7
Additional Analyses

Panel A: Results of Meeting or Just Beating Analysts' Forecasts Based on the Ratio of Stock and Option Grants to Total Compensation

This panel reports the results from regressing the probability of earnings surprises in a certain range on equity incentives and controls:

$$Prob(Surprise_{i,t} \in \Phi_k) = logit(\gamma_{0,k} + \beta_k \mathrm{Eq\_Incent}_{i,t} + \gamma_k Controls_{i,t} + \zeta_{i,t}),$$

where equity incentives include a stock and option grant ratio and ownership, and controls are those included in Equation (2). The table reports the change in the odds due to a change of one standard deviation of a variable and the p-value (in parentheses) of the corresponding coefficient in the logit regressions. The stock and option grant ratio is defined as the ratio of stock grants and option grants (both in dollars) to the total compensation, which includes salary, bonus, stock grants, option grants, long-term incentive payouts, and other compensation. See Appendix A for the measurement of other variables. Year dummies are included to control for year fixed effects. The regressions are based on 4,274 firm-years with all required data in the period 1993–2000.

	Negative Earnings Surprises $\phi = (-\infty, 0)$ (1)		**Meeting or Just Beating Analysts' Forecasts $\phi = [\$0, \$0.01]$ (2)**	
	Predicted Signs	**Change in the Odds (p-value)**	**Predicted Signs**	**Change in the Odds (p-value)**
Equity incentives				
Stock and option grant ratio	−	−0.103	+	0.105
		(0.004)		(0.002)
Ownership	−	−0.203	+	0.297
		(0.001)		(0.001)
Control variables		Yes		Yes
Pseudo R^2		0.195		0.180
Likelihood ratio		603		612
(p-value)		(0.001)		(0.001)

(continued on next page)

TABLE 7 (continued)

Panel B: Analyses of Future Net Sales Conditional on Future Realized Stock Returns

This panel reports results from regressing future net sales (%) on equity incentives (%), a high return dummy variable (*High_ret*), the interactions between equity incentives and the high return dummy, and control variables:

$$\begin{aligned} NetSale_{i,t+j} = {} & \gamma_0 + \beta\, Eq\_Incent_{i,t} + \delta\, Eq\_Incent_{i,t} \times High\_ret_{i,t+j} + \chi\, High\_ret_{i,t+j} \\ & + \gamma_1 Size_{i,t} + \gamma_2 Growth_{i,t} + \gamma_3 Ret_{i,t} + \gamma_4 \Delta OCF_{i,t} + \zeta_{i,t+j}, \end{aligned}$$

where $j = 1$ or 2, denotes the *j*th six-month period after the earnings announcement. $High\_ret_{i,t+j}$ is 1 when the return in the period $t+j$ is in the top quartile. See Appendix A for the measurement of other variables. The regressions are estimated on all firm-years with data available in the period 1993–2000. Year dummies are included to control for year fixed effects. The p-values in parentheses are based on one-sided t-tests.

	Predicted Signs	1st Six-Month	2nd Six-Month
Equity incentives			
Unexercisable options	+	0.056	0.017
		(0.003)	(0.052)
Ownership	+	0.012	0.008
		(0.001)	(0.001)
Exercisable options	+	0.001	0.006
		(0.476)	(0.374)
Equity incentives × *High_ret*			
Unexercisable options × *High_ret*	+	0.080	0.230
		(0.017)	(0.001)
Ownership × *High_ret*	+	0.016	0.023
		(0.001)	(0.001)
Exercisable options × *High_ret*	+	−0.025	0.046
		(0.826)	(0.072)
High_ret	+	0.032	−0.031
		(0.337)	(0.565)
Control variables		Yes	Yes
n		2,723	2,316
Adj. R^2		0.129	0.133

Results for Subperiods

To ensure that boom markets in late 1990s do not drive our results, we split the sample into two subsamples: 1993–1996 and 1997–2000. The results in these two subsamples are similar except that the impact of unexercisable options on net sales and earnings management, while significant in both subsamples, is slightly stronger in later years than in earlier years.

Alternative Explanations

One might argue that our results on meeting or beating analysts' forecasts are consistent with better firm performance attendant to the incentive alignment benefit of equity incentives. This alternative explanation is unlikely to drive our results for at least two reasons. First, financial analysts are generally regarded, and empirically confirmed, as sophisticated information intermediaries; thus, analysts arguably have incorporated into their forecasts the performance improvement due to equity incentives. If so, then forecast errors should not be correlated with equity incentives. Second, better firm performance arising from incentive alignment benefits cannot explain the evidence that firms with high equity incentives are *less* likely to have large positive earnings surprises. Being less likely to have large positive earnings surprises is not a signal of good performance, but is consistent with our earnings smoothing argument.

Another alternative explanation relates to the possibility that our results are driven by equity incentives proxying for firm characteristics, such as growth, risk, or earnings volatility. In the main and sensitivity analyses, we explicitly control for these factors. Also, untabulated results indicate that firms with volatile earnings (high *CV_AF*) are less likely to meet or just beat analysts' forecasts and more likely to have large positive earnings surprises, but the results on equity incentives are just the opposite. Furthermore, while both exercisable and unexercisable options are similarly related to firm characteristics (as shown in Appendix B), exercisable options, unlike unexercisable options as explained above, are not associated with future trading or earnings management measures. Such evidence suggests that omitted correlated variables are unlikely to drive our results.

Finally, to control for industry characteristics (e.g., performance volatility and earnings informativeness) that might affect the level of equity incentives and the probability of meeting or just beating analysts' forecasts, we replicate all analyses using an industry-demeaned equity incentive measure. To this end, we combine unexercisable options, weighted by their hedge ratios, with stock ownership and then deduct from this measure its mean in the industry-year level (industry is defined based on three-digit SICs.) The inferences based on this measure are essentially the same. While we cannot completely rule out the possibility that these alternative explanations contribute to the reported results, we believe that our results are more consistent with earnings management induced by equity incentives rather than with these alternative explanations.

VI. SUMMARY AND CONCLUSIONS

In this paper, we examine the relation of equity incentives—broadly defined as stock-based compensation and stock ownership—with managers' future trading and earnings management. We first study the link between equity incentives and managers' future trading. Based on all firm-years with available data over the 1993–2000 period, we find that high equity incentive managers are more likely to sell shares in the year after earnings announcements. This selling behavior is consistent with managers' diversifying the increased risk arising from equity incentives, as argued in the paper, and/or with managers exploiting inside information on the pricing of their companies' stocks. As stated by Michael Jensen, "Once a firm's shares become overvalued, it is in managers' interests to keep them that way, or to encourage even more overvaluation, in the hope of cashing out before the bubble bursts" (*The Economist* 2002). Thus, managers with high equity incentives could benefit from earnings management with the objective of keeping stock prices high and increasing the value of the shares to be sold in the future.

We then test whether high equity incentive managers engage in earnings management by examining whether these managers are more likely to report earnings that meet or just beat analysts' forecasts. The analyses indicate a significantly higher incidence of meeting or just beating analysts' forecasts for managers with high equity incentives. Additional analyses document that managers with high equity incentives, especially those with consistently high equity incentives, are less likely to report large positive earnings surprises, consistent with earnings smoothing.

Overall, our results suggest that stock-based compensation and ownership can lead to incentives for earnings management. If earnings management can increase short-term stock prices, managers can benefit from doing so by increasing the value of the shares they are going to sell. Results from additional analyses indicate that high equity incentive managers can even gain more by selling more shares after reporting earnings that meet or beat analysts' forecasts or after recognizing income-increasing abnormal accruals. Consequently, such earnings management behavior likely increases managers' wealth at the expense of outside shareholders. Thus, this paper extends research on compensation-related opportunistic behavior, such as Aboody and Kasznik (2000).

Our findings should be of interest to boards of directors contemplating compensation contracts for managers. While stock-based compensation and ownership can yield positive incentive effects, they can also have side effects, such as increased earnings management. An optimal compensation contract should reflect this trade-off. Our results also have implications for financial reporting regulators, who have criticized managers for focusing too much on short-term performance targets, potentially to the detriment of earnings quality (Levitt 1998). Given the increased importance of stock-based compensation as a form of executive compensation and managerial ownership as an incentive alignment mechanism, our analyses indicate that regulatory responses to address opportunistic earnings management should consider incentives arising from stock-based compensation and ownership.

APPENDIX A
Variable Measurement

Variables **Measurement**

Panel A: Equity Incentives

$Eq\_Incent_{i,t}$ = equity incentives are measured in shares during or at the end of fiscal year t, scaled by the number of outstanding shares; this measure includes five components: option grants, unexercisable options (excluding option grants), exercisable options, restricted stock grants, and ownership. Unless noted, restricted stock grants are combined with ownership and option grants are combined with other unexercisable options.

Panel B: Future Net Sales

$NetSale_{i,t+j}$ = CEO's net sales (in dollars) in the jth six-month period after the earnings announcement for fiscal year t, scaled by the market value at the end of fiscal year t, where $j = 1$ and 2, and net sales = open market sales − (open market purchases + options exercised).

Panel C: Earnings Surprises

$Surprise_{i,t}$ = difference between actual earnings and the latest consensus forecast before earnings announcements, both of which are measured on a per-share basis and are collected from First Call; it is rounded to the nearest cent; and

$D\_MB_{i,t}$ = meeting or just beating dummy variable, 1 if earnings surprises are zero or one cent, and 0 otherwise.

Panel D: Abnormal Accruals

$Abnormal\ Accruals_{i,t}$ = abnormal accruals estimated from the modified cross-sectional Jones (1991) model (see Panel A of Table 6 for the estimation details).

Panel E: Factors Affecting Future Net Sales

$Size_{i,t}$ = natural logarithm of market value (in million dollars) at the end of fiscal year t;

$Growth_{i,t}$ = the book-to-market ratio at the end of fiscal year t;

$Ret_{i,t}$ = buy and hold raw return in the 12 months prior to the earnings announcement for fiscal year t; and

$\Delta OCF_{i,t}$ = change in operating cash flow in year t, scaled by total assets at the beginning of year t.

Panel F: Factors Affecting Earnings Surprises

$Size_{i,t}$ and $Growth_{i,t-1}$ as measured in Panel E.

$Sales\_Growth_{i,t}$ = sales for fiscal year t divided by sales for fiscal year $t-1$;

$NOA_{i,t-1}$ = net operating assets (i.e., shareholders' equity minus cash and marketable securities, plus total debt) at the end of fiscal year $t-1$, scaled by sales for fiscal year $t-1$;

$Shares_{i,t}$ = number of common shares outstanding at the end of fiscal year t;

$Litigation_{i,t}$ = 1 if the firm is in one of the following industries: pharmaceutical/biotechnology (SIC codes 2833–2836, 8731–8734), computer (3570–3577, 7370–7374), electronics (3600–3674), or retail (5200–5961), and 0 otherwise;

$ImplicitClaim_{i,t}$ = proxied by labor intensity, calculated as 1 minus the ratio of gross PPE to total assets at the end of fiscal year t;

$Num\_Estimate_{i,t}$ = number of analysts whose forecasts are included in the consensus forecast used to calculate earnings surprise;

$CV\_AF_{i,t}$ = coefficient of variation (standard deviation scaled by the mean) of the consensus forecast used to calculate earnings surprise; and

$Down\_Rev_{i,t}$ = 1 if at least one of the firm's analysts revised his or her forecast downward in the three months prior to the earnings announcement for fiscal year t, and 0 otherwise.

Panel G: Factors Affecting Abnormal Accruals

$Size_{i,t}$ and $Growth_{i,t}$ as measured in Panel E; $NOA_{i,t}$, $Litigation_{i,t}$, and $ImplicitClaim_{i,t}$ as measured in Panel F.

$Leverage_{i,t}$ = long-term debt divided by total assets, both at the end of fiscal year t;

$Bonus_{i,t}$ = 1 for firm-years whose CEOs have nonzero bonus, and 0 otherwise; and

$Risk_{i,t}$ = market beta in fiscal year t estimated from weekly returns.

APPENDIX B
Correlations between Variables

This appendix reports Spearman correlations between key variables used in various analyses. Correlations between all variables, except for those noted below, are based on the sample used in earnings surprise analyses, 4,301 firm-years. Correlations between net sales or stock returns and other variables are based on the sample used in future trading analyses, 2,723 firm-years. Correlations between leverage or risk and other variables are based on the sample used in abnormal accrual analyses, 6,387 firm-years. If two variables are not included in the same analysis, then we do not report their correlation and use NA to denote such occasions. If all correlations in a column or a row are not reported, then this column or this row is dropped from the table. See Appendix A for variable measurement. All correlations, except those denoted with #, are significant at the 0.05 level or lower.

	Net Sales[a]	*D_MB*	Abnormal Accruals	Unexercisable Options	Ownership	Exercisable Options	Size	Growth	Net Operating Assets	Litigation	Implicit Claim	CV_AF	Leverage
Abnormal accruals	0.005#	0.035											
Unexercisable options	0.073	0.037	0.021#										
Ownership	0.193	0.049	0.030	0.050									
Exercisable options	−0.040#	−0.027#	0.003#	0.472	−0.011#								
Size	0.084	0.140	−0.003#	−0.209	−0.343	−0.243							
Growth	−0.282	−0.194	−0.070	0.026	−0.025	0.108	−0.407						
Stock returns	0.269	NA	−0.019#	0.071	0.019#	−0.003#	0.209	−0.368					
Net operating assets	NA	−0.030	−0.012#	−0.027	−0.132	0.036	0.051	0.218					
Litigation	NA	0.120	−0.008#	0.068	0.097	0.037	0.057	−0.194	−0.201				
Implicit Claim	NA	0.084	0.017#	0.138	0.106	0.088	−0.025	−0.153	−0.211	0.135			
CV_AF	NA	−0.203	−0.151	−0.005#	0.093	0.009#	−0.196	0.177	0.073	0.070	−0.058		
Down_Rev	NA	−0.001#	0.004#	−0.063	−0.100	−0.048	0.070	0.027	0.104	−0.047	−0.071	0.081	
Leverage	NA	−0.036	0.032	−0.148	0.090	0.002#	0.194	−0.036	0.353	−0.261	−0.271	NA	
Risk	NA	−0.022#	0.066	0.035	−0.010#	0.141	−0.199	−0.022	−0.067	0.238	0.119	NA	−0.133

---

[a] For the sake of brevity, we only report the correlations between net sales in the 1st six-month period after earnings announcements and other variables. The correlations between net sales in the 2nd six-month period and other variables are similar.

REFERENCES

Abarbanell, J. S., and V. L. Bernard. 1992. Test of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *The Journal of Finance* 47 (3): 1182–1207.

Aboody, D., and R. Kasznik. 2000. CEO stock option awards and timing of corporate voluntary disclosure. *Journal of Accounting and Economics* 29 (1): 73–100.

Ali, A., and S. Kallapur. 2001. Securities price consequences of the Private Securities Litigation Reform Act of 1995 and related events. *The Accounting Review* 76 (3): 431–460.

Baber, W. R., and S. H. Kang. 2002. The impact of split adjusting and rounding on analysts' forecast error calculation. *Accounting Horizons* 16 (4): 277–289.

Baker, G. P., and B. J. Hall. 2004. CEO incentives and firm size. *Journal of Labor Economics* 22 (4): 767–798.

Barton, J., and P. J. Simko. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* 77 (Supplement): 1–27.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (2): 173–204.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley & Sons.

Beneish, M. 1999. Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74 (4): 425–457.

———, and M. Vargus. 2002. Insider trading, earnings quality, and accrual mispricing. *The Accounting Review* 77 (4): 755–791.

Bernard, V. L., and D. J. Skinner. 1996. What motivates managers' discretionary accruals? *Journal of Accounting and Economics* 22 (1–3): 313–325.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3): 637–654.

Bowen, R. M., L. DuCharme, and D. Shores. 1995. Stakeholders' implicit claims and accounting method choice. *Journal of Accounting and Economics* 20 (3): 255–295.

Brown, L. D., and K. J. Kim. 1991. Timely aggregate analysts' forecasts as better proxies for market earnings expectations. *Journal of Accounting Research* 29 (2): 382–385.

Burgstahler, D. C., and M. J. Eames. 2003. Earnings management to avoid losses and earnings decreases: Are analysts fooled? *Contemporary Accounting Research* 20 (2): 253–294.

Conference Board. 2002. Executive compensation: Principles, recommendations and specific best practice suggestions. In *Commission on Public Trust and Private Enterprise: Findings and Recommendations*, September 17, 2002. New York, NY: Conference Board.

Core, J. E., and W. R. Guay. 1999. The use of equity grants to manage optimal equity incentive levels. *Journal of Accounting and Economics* 28 (2): 151–184.

———, and D. F. Larcker. 2002. Performance consequences of mandatory increases in executive stock ownership. *Journal of Financial Economics* 64 (3): 317–340.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–250.

DeFond, M. L., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals: Accounting choice in troubled companies. *Journal of Accounting and Economics* 17 (1–2): 145–176.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (1): 1–35.

Economist, The. 2002. How to pay bosses. (November 14).

Financial Accounting Standards Board (FASB). 2004. *Share-Based Payment, an Amendment of FASB Statements No. 123 and 95*. Proposed Statement of Financial Accounting Standards. File Ref. 1102-100, March 31, 2004. Stamford, CT: FASB.

Fields, T. D., T. Z. Lys, and L. Vincent. 2001. Empirical research on accounting choices. *Journal of Accounting and Economics* 31 (1–3): 255–307.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (2): 137–164.

Hanlon, M., S. Rajgopal, and T. Shevlin. 2003. Are executive stock options associated with future earnings. *Journal of Accounting and Economics* 36 (1–3): 3–43.

Healy, P. M. 1985. The impact of bonus schemes on the selection of accounting principles. *Journal of Accounting and Economics* 7 (3): 85–107.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics* 3 (4): 305–360.

———, and K. J. Murphy. 1990. Performance pay and top-management incentives. *Journal of Political Economy* 98 (2): 225–264.

Jones, J. 1991. Earnings management during import relief investigation. *Journal of Accounting Research* 29 (2): 193–228.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37 (1): 57–82.

———, and M. F. McNichols. 2002. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40 (3): 727–759.

Ke, B. 2001. Why do CEOs of publicly traded firms prefer reporting small earning increases and long strings of consecutive earnings increases. Working paper, Pennsylvania State University.

———, S. Huddart, and K. Petroni. 2003. What insiders know about future earnings and how they use it: Evidence from insider trades. *Journal of Accounting and Economics* 35 (3): 315–346.

Knowledge at Wharton. 2003. Stock options: The end of the affair. Available at: http://knowledge.wharton.upenn.edu.

Lakonishok, J., and I. Lee. 2001. Are insider trades informative? *Review of Financial Studies* 14 (1): 79–111.

Lambert, R. A., and D. F. Larcker. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 85–129.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, September 28, New York, NY.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

McNichols, M. F., and G. P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

McVay, S., V. Nagar, and W. Tang. 2004. Stock incentives to meet earnings thresholds. Working paper, University of Michigan.

Morck, R., A. Shleifer, and R. W. Vishny. 1988. Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics* 20 (1–2): 293–315.

Morgenson, G. 2003. The rules on bosses' pay seem written with pencil. *New York Times* (May 25).

Murphy, K. J. 1999. Executive compensation. In *Handbook of Labor Economics*, edited by O. Ashenfleters and D. Card. Amsterdam, The Netherlands: North Holland.

Noe, C. F. 1999. Voluntary disclosures and insider transactions. *Journal of Accounting and Economics* 27 (3): 305–326.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10 (1): 53–83.

Ofek, E., and D. Yermack. 2000. Taking stock: Equity-based compensation and the evolution of managerial ownership. *The Journal of Finance* 55 (3): 1367–1384.

Payne, J. L., and S. W. G. Robb. 2000. Earnings management: The effect of ex-ante earnings expectations. *Journal of Accounting, Auditing & Finance* 15 (4): 371–392.

Penman, S. H. 1982. Insider trading and the dissemination of firms' forecast information. *The Journal of Business* 55 (4): 479–503.

Rozeff, M. S., and M. A. Zaman. 1998. Overreaction and insider trading: Evidence from growth and value portfolios. *The Journal of Finance* 53 (2): 701–716.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (4): 91–102.

Seyhun, H. N. 1998. *Investment Intelligence from Insider Trading*. Cambridge, MA: The MIT Press.

Sivakumar, K., and G. Waymire. 1994. Insider trading following material news events: Evidence from earnings. *Financial Management* 23 (1): 23–32.

Skinner, D. J., and L. Myers. 2000. Earnings momentum and earnings management. Working paper, University of Michigan.

———, and R. G. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (2–3): 289–312.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Stein, J. C. 1989. Efficient capital markets, inefficient firms: A model of myopic corporation behavior. *Quarterly Journal of Economics* 104 (4): 655–669.

Summers, S., and J. Sweeney. 1998. Fraudulently misstated financial statements and insider trading: An empirical analysis. *The Accounting Review* 73 (1): 131–146.

Sweeney, A. P. 1994. Debt-covenant violations and managers' accounting response. *Journal of Accounting and Economics* 17 (3): 281–308.

Thomas, J., and H. Zhang. 2002. Inventory changes and future returns. *Review of Accounting Studies* 7 (2–3): 163–187.

Warfield, T. D., J. J. Wild, and K. L. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20 (1): 61–91.

Watts, R. L., and J. L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (3): 357–373.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 477–500

The Effect of Control Systems on Trust and Cooperation in Collaborative Environments

Angela L. Coletti
The University of Texas at Austin

Karen L. Sedatole
The University of Texas at Austin

Kristy L. Towry
Emory University

ABSTRACT: Because of conflicting incentives among participants, collaborations (e.g., strategic alliances, joint ventures, and work teams) present a significant control challenge to managerial accountants. On the one hand, formal controls such as sanctioning and monitoring systems improve cooperation by reducing the *incentives* for opportunistic behavior. On the other hand, prior research finds that the mere presence of a control system causes decision makers to view the collaborative setting as non-cooperative, and other collaborators as untrustworthy. In this paper, we conduct two experiments in which participants act as business collaborators. Through these experiments, we examine the effects of control on trust and cooperation in collaborative settings. Specifically, we posit and provide evidence that a strong control system can enhance the level of trust among collaborators. The mediating role of control-induced cooperation provides the mechanism by which control systems can increase trust in collaborative environments. Furthermore, we show that this increased trust has a positive effect on the subsequent level of cooperation among collaborators. Taken together, the results suggest an increasing marginal benefit of control system strength arising from the trust that control-induced cooperation engenders. The implication is that firms will choose to implement a stronger control system than previous research would seem to suggest.

This paper has benefited from useful discussions with Shannon Anderson, Urton Anderson, Harry Evans, R. Lynn Hannan, Vicky Hoffman, Bruce Johnson, Steve Kachelmeier, William Kinney, Lisa Koonce, Marlys Lipe (editor), Molly Mercer, Don Moser, Bill Rankin, Casey Rowe, Steve Salterio, Janet Samuels, Rick Tubbs, Wim Van der Stede, two anonymous reviewers, workshop participants at Arizona State University, The University of Iowa, University of Pittsburgh, and The University of Texas at Austin, and participants at the AAA MAS 2004 Midyear Meeting and the 2004 Southeast Summer Accounting Research Colloquium. We gratefully acknowledge Pritesh Ghaghada for his programming support, and Romana Autrey, Kirill Novoselov, and Jessen Hobson for their help in administering the experiments. We appreciate the financial support of the McCombs School of Business Center for Business Measurement and Assurance Services, and the Eugene and Dora Bonham Memorial Fund.

Editor's note: This paper was accepted by Marlys Gascho Lipe, Editor.

Submitted August 2003
Accepted July 2004

Keywords: *collaboration; strategic alliance; joint venture; teams; control system; Attribution Theory; framing; trust; cooperation.*

I. INTRODUCTION

Over the past several decades, collaborative enterprise has emerged as an important business trend. Popular forms of collaboration include teams, through which companies seek gains from the synergy of employees working together toward common goals, as well as strategic alliances, joint ventures, and industry consortia, through which companies reach outside their boundaries to gain a competitive advantage. This new collaborative environment presents a managerial control challenge to firms as they attempt to identify optimal levels of control mechanisms while simultaneously considering the impact of such mechanisms on interpersonal trust. Prior research finds that trust is a primary determinant of collaboration success (e.g., Spekman et al. 2001; Zaheer et al. 1998; Zaheer and Venkatraman 1995). However, some scholars suggest that control systems reduce trust among collaborators. Several theories have been used to make this argument. For example, some scholars argue that the mere presence of a control system changes the way decision makers mentally frame the situation, causing them to perceive other collaborators as less trustworthy (e.g., Tenbrunsel and Messick 1999). Others suggest that cooperation in the presence of a control system is attributed to the constraints imposed by the control system rather than the cooperator's inherent trustworthiness, thus inhibiting the development of trust (e.g., Malhotra and Murnighan 2002).

In this paper we challenge the view that control systems reduce trust. We argue that prior studies documenting the negative effects of control systems are limited because they fail to incorporate two common features of control systems. First, prior studies do not allow for the increased cooperation that is likely to arise from the economic incentives inherent in control systems. For example, Malhotra and Murnighan (2002) compare the level of trust that develops when control systems are present versus absent. In that study, participants believed they were interacting with other participants via a computer network. However, these other participants were actually computers pre-programmed to always cooperate. Thus, Malhotra and Murnighan (2002) hold the level of collaborative cooperation constant across conditions. Second, prior studies do not allow for feedback regarding levels of cooperation (e.g., performance reports). Tenbrunsel and Messick (1999), for example, simply do not provide participants with feedback on the level of collaborative cooperation. Our study, however, allows for control-induced cooperative behavior that is observed by participants. Specifically, we posit and provide evidence that a control system can actually enhance the level of trust among collaborators, provided the control system is strong enough to induce cooperation and that this cooperation is observed by the collaborators. Further, we show that this increased trust has a positive effect on the future level of cooperation among collaborators. Taken together, the results suggest an increasing marginal benefit of control system strength arising from the trust that control-induced cooperation engenders. The implication is that firms will choose to implement a stronger control system than previous research would suggest.

This study is important to managers who form collaborations and to managers, managerial accountants, and internal auditors charged with establishing managerial control systems. While firms have eagerly embraced collaboration as a key strategy, many types of collaborations are notoriously unstable, with a number of experts attributing their high failure rate to a lack of trust. This study suggests that one approach to increasing trust is to strengthen the control systems used to govern collaborative agreements, both in terms of their incentives and their feedback. That is, through increased monitoring, sanctioning, and rewarding, firms can induce higher levels of cooperation. The control-induced cooperation, when observed by collaborators via feedback mechanisms such as performance reports, will engender trust, thereby reinforcing the positive effects of the control mechanisms. Thus, this paper answers recent calls for research to provide "insights regarding whether explicit contracts based on managerial accounting information foster or destroy reciprocity and cooperation" (Sprinkle 2003, 296).

We conduct two experiments to explore the effects of control on trust and cooperation. The first experiment is conducted using the traditional approach of psychology-based experimental work, whereas the second uses an approach more representative of the experimental economics literature. We exploit the advantages of each method in order to "triangulate" the results and develop a more comprehensive understanding of the complex relations among control, trust, and cooperation. Thus, this study illustrates the potential for research that combines economics and psychology-based theory to provide insights into accounting issues (Haynes and Kachelmeier 1998; Merchant et al. 2003; Moser 1998; Waller 1995, 2001).

In both experiments, participants assume the role of a manager involved in a collaboration. In Experiment 1, we manipulate the presence or absence of a control system and randomly assign participants such that inherent trustworthiness is not expected to differ across conditions. We find that participants cooperate more in the *control system* condition, suggesting that the control system manipulation provides incentives strong enough to induce cooperation. Importantly, a second set of observer participants rate the participants assigned to the *control system* condition as more trustworthy than those assigned to the *no control system* condition. Since the inherent trustworthiness is not expected to differ across conditions, it appears that the observer participants attribute the control-induced cooperation to the inherent trustworthiness of the managers. Thus, we find that the presence of a control system enhances the level of trust in a collaborative setting, and that observed cooperation perfectly mediates this effect.

We corroborate and extend these findings in the second experiment, which, unlike Experiment 1, allows us to investigate the effect of trust on subsequent cooperative behavior. We again manipulate the presence or absence of a control system, randomly assigning participants such that inherent trustworthiness is not expected to differ across conditions. Participants in this experiment make decisions under explicit incentives and salient cash payments. Over 20 periods of play, we again find that cooperation is greater for those participants operating under a control system than for those operating without one. After 20 periods, we remove the control system, such that, although the conditions vary in terms of the interactive histories, the incentives *going forward* are equivalent across the two experimental conditions. We find that trust is greater for those participants who had been operating under a control system than for those who had not. We again document that observed cooperation perfectly mediates the effect of the control system on trust. Further, we find that participants previously operating under a control system continue to be more cooperative than those previously operating without one.

The remainder of the paper is organized as follows: In Section II, we review the relevant literature and develop the hypotheses. In Section III, we describe the experimental methodology and results for each experiment. Section IV concludes and discusses the implications of the study.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

The Growing Use of Collaborations

We define a collaboration as two (or more) parties involved in some type of joint production, such that each party's individual output cannot be directly observed. Our definition applies generally to many different types of collaborations that are being increasingly used to establish competitive advantage. For example, in a survey of 700 American workers by Dale Carnegie Training, 90 percent of the respondents reported spending some time working in teams (Allerton 1996). Firms also increasingly engage in alliances with other firms (e.g., licensing agreements, joint ventures, industry consortia, and strategic supply chains). (See Anderson and Sedatole [2003] and Birnberg [1998] for a discussion of collaboration types and their inherent control problems.) According to a recent survey, the average firm in a group of 90 multinationals formed 117 alliances over a three-year period (Gordon et al. 2000). In fact, many now view relational capital—the ability to attract and interact effectively with alliance partners—as a core competency (Schifrin 2001).

Risk and Control in Collaborations

While collaborations undoubtedly provide many advantages, they are also quite unstable (Church 1996; Das and Teng 2000; Dumaine 1994; Parkhe 1993). Collaborations are vulnerable to failure because they are exposed to both "performance risk" and "relational risk" (e.g., Das and Teng 1996, 1999, 2001). Performance risk is the probability that collaboration objectives will not be achieved despite the full cooperation of the partners. This type of risk comes primarily from market forces such as competition and demand fluctuations and is, therefore, not unique to collaborations. Relational risk, on the other hand, is the probability that collaboration objectives will not be achieved because of a lack of cooperation. This risk is unique to collaborations and is, therefore, the focus in this paper.

High levels of relational risk arise because of the social dilemma nature of collaborations. *Social dilemmas* are situations where opportunistic behavior leads to suboptimal outcomes but cooperative behavior is not individually rational (Dawes 1980). Collaborations are formed because of the belief that there are gains from cooperation and information sharing. However, the difficulty of measuring individual contributions to collaborative output creates an incentive for opportunistic behavior, as collaborators are tempted to free-ride and withhold information. Further, if all collaborators behave in this manner, then the collaboration itself is destined to fail.

Management accountants can play an important role in mitigating relational risk through the design and implementation of formal control mechanisms, such as systems of monitoring, sanctioning, and rewarding (Seal et al. 1999). Control, broadly defined, comprises the various policies and procedures that firms use to mitigate different types of risk. Controls mitigate relational risk by changing the *incentives* for opportunistic behavior. That is, periodic monitoring increases the probability that opportunistic behavior will be detected, and sanctioning (rewarding) systems impose penalties (bonuses) on collaborators who (do not) engage in such behavior. Prior empirical research documents the use of formal control mechanisms to reduce relational risk in collaborations. Ittner et al. (1999), for example, document benefits of increased selection and monitoring practices (e.g., frequent meetings and supplier certification programs) for supplier partnerships (see also, Groot and Merchant 2000; Clement 1997). Others have documented the benefits of performance measurement (e.g., Meyer 1994) and surveillance (e.g., Sewell 1998) in collaborative settings.

From a theoretical perspective, agency theory provides insights into the optimal level of control for mitigating relational risk in collaborative settings.[1] Using a similar approach, Figure 1, Panel A illustrates the fundamental trade-off between the cost of control and the costs incurred when collaborators behave opportunistically. Specifically, we assume that the expected cost of opportunistic behavior (i.e., relational risk) decreases with control system strength. Offsetting this benefit is the cost of implementing and maintaining the control system, which increases with control system strength. As the figure illustrates, there is an implied optimal level of control for which total costs are minimized.[2] Importantly, this optimal control level likely results in something less than full cooperation. In other words, the firm must assume some relational risk because complete contracts (i.e., providing full cooperation) are prohibitively expensive, largely due to the performance ambiguity inherent in collaborations.

As implied by Figure 1, Panel A, in choosing the level of control, it is important for firms to understand the relation between the strength of the control system and total cost. Below, we explore how this relation changes when *trust* is added to the equation (Chenhall and Langfield-Smith 2003; Spekman et al. 2001; Zaheer et al. 1998; Zaheer and Venkatraman 1995).

Control and Trust

We define *trustworthiness* as an innate personal characteristic reflecting one's preference for upholding some social norm of behavior, regardless of economic incentives.[3] In contrast, we consider *trust* to be one's perception of another's trustworthiness. Because collaborations are formed specifically to capture gains from synergy, it is reasonable to assume that cooperation is viewed as a relevant social norm in these settings. Thus, trust is conceptualized as the perceived likelihood that another person will cooperate, absent any economic incentives to do so (i.e., "I trust you to do the right thing, because you are a trustworthy person"). It is important to note that our definition differs from a game-theory perspective (e.g., Williamson 1993), which assumes that trust occurs when the economic incentives favor cooperative behavior (i.e., "I trust you to do the right thing, because it is not in your best interest to do the wrong thing"). Throughout this paper, the term *trustworthiness* refers to an innate personal characteristic, while *trust* refers to an individual's perception of another's trustworthiness.[4]

Prior research suggests that control systems can have unintended negative effects on trust, a finding of particular concern because of trust's critical role as a determinant of collaborative success (Das and Teng 1998; Jarvenpaa et al. 1998; Kale et al. 2000; Kramer and Tyler 1996; Landry 1998; McAllister 1995; Spagnolo 1999). If this claim is correct, then this presents a conundrum to the designers of managerial control systems, who must choose the optimal level of control, considering both the benefits (through reduced incentives for opportunistic behavior) and the costs (including a degradation of trust).

---

[1] Specifically, this research shows that firms may make cost-saving reductions in explicit incentives (i.e., weaken the control system) in contracting with employees because of the supporting implicit incentives (e.g., peer pressure, mutual monitoring) provided by team members (Arya et al. 1997; Barron and Gjerde 1997; Che and Yoo 2001). Balakrishnan et al. (1998) further show a trade-off between the cost and benefits (team performance improvements) in choosing the strength of the control system (i.e., the informativeness of an audit).

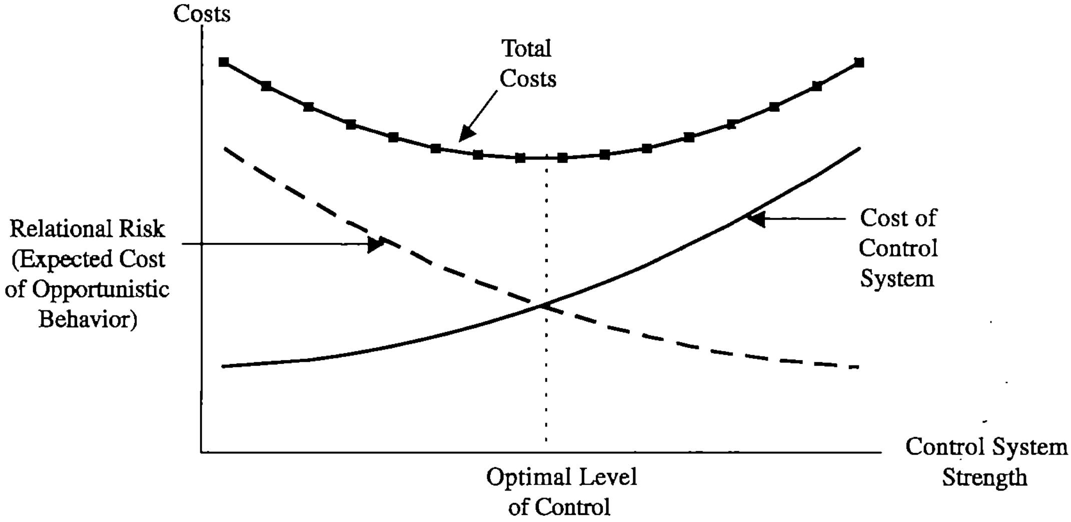
[2] We do not claim to have operationalized an optimal control system, because to do so would require a specification of control system cost, which is outside the scope of this study. Rather, we examine how incorporating trust into the model shifts the optimal level of control.

[3] Relevant social norms may include honesty, fairness, cooperation, commitment, reciprocity, accountability, etc.

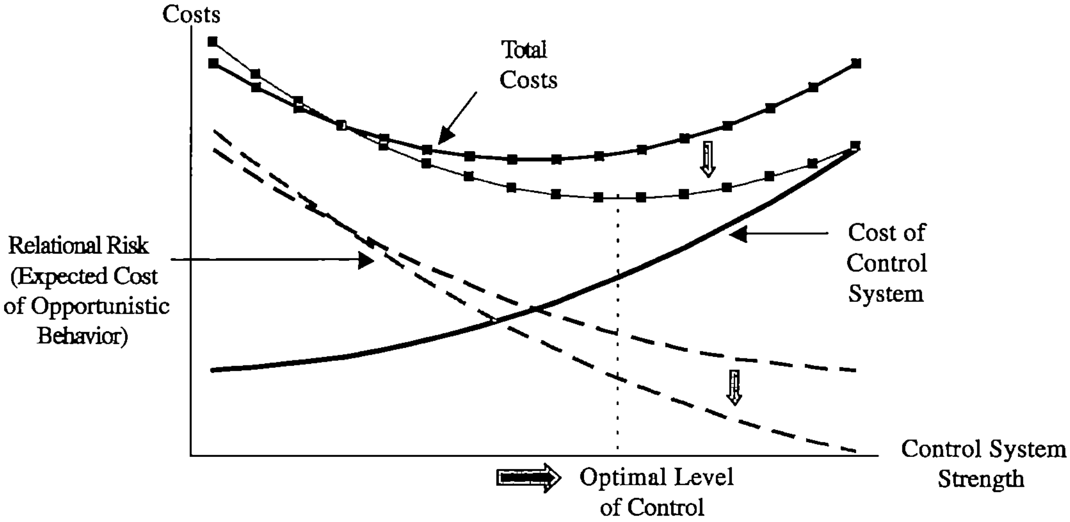
[4] See Kramer (1999) for other conceptions of trust in an organizational setting.

FIGURE 1
Control Costs, Relational Risk, and the Optimal Level of Control

Panel A: Fundamental Trade-Off between Control Costs and Relational Risk



Panel B: The Effect of Trust on the Optimal Level of Control



Adapted from Juran (1974).

Two main bodies of theory have been proposed to explain why control systems may reduce trust and cooperation in collaborative settings. The first body of theory is related to *framing*. According to this theory, decision making in social situations begins with a categorization process. That is, the decision maker asks, "What kind of situation is this?" (March and Heath 1994, 58). The answer to this question depends on environmental cues.

The resulting categorization, or *decision frame*, has a powerful influence on the decision maker, affecting not only the perceptions of what norms are applicable, but also the expectations of others (Messick 1999).

Firms form collaborations specifically to capture gains from synergy and cooperation. Therefore, the collaborative setting itself serves as an environmental cue, and in the base case, decision makers are likely to frame the situation as inherently cooperative, and other collaborators as inherently trustworthy. Tenbrunsel and Messick (1999) suggest, however, that the introduction of a control system may serve as another cue, affecting how the situation is categorized. Specifically, they examine social dilemmas in a collaborative setting, and find that a control system invokes a business (as opposed to ethical) frame. Consequently, with a control system, collaborators focus on a logical and rational evaluation of the costs and benefits of cooperation, rather than on social norms of behavior. This frame increases the salience of prospects for opportunistic behavior by both the decision maker and others. In fact, Tenbrunsel and Messick (1999) provide evidence that both cooperation and trust (i.e., expectations of cooperation) are lower under a weak control system than under no control system.[5] They conclude that control systems may "have a negative influence on perceptions, perhaps promoting a distrust of coworkers" (Tenbrunsel and Messick 1999, 693).

Second, Attribution Theory (Birnberg et al. 1977; Jones 1990; Kelley 1967) has been used to examine the effects of control systems on perceived trustworthiness and trusting behavior (e.g., Kruglanski 1970; Malhotra and Murnighan 2002; Strickland 1958). Attribution Theory comprises a wide body of literature aimed at understanding how people make causal attributions regarding, for example, the behavior of others. Generally speaking, behaviors can be attributed to either dispositional or situational factors. In the present context, when a collaborator cooperates, other collaborators may attribute the cooperation either to the person's innate trustworthiness (a dispositional characteristic) or to the fact that there is a control system in place (a situational characteristic). For example, Malhotra and Murnighan (2002) find that cooperation in the presence of a control system (i.e., binding contract) is attributed at least partially to the constraints imposed by the control system. On the other hand, cooperation in the absence of a control system (i.e., nonbinding contract) is perceived to be a reflection of a person's trustworthiness. Malhotra and Murnighan (2002) conclude that control systems inhibit the development of trust.

Limitations of Prior Research

Prior studies fail to incorporate two common features of control systems: (1) increased cooperation induced by the economic incentives control systems provide, and (2) feedback regarding that cooperation (e.g., performance reports). Because of these two exclusions, prior studies inhibit the positive effects of control systems. Some studies hold the level of collaborative cooperation constant. For example, participants in the Malhotra and Murnighan (2002) study believed they were interacting with other participants via a computer network. However, these other participants were actually computers pre-programmed to always cooperate. Thus, this study compares how trust develops when individuals observe others cooperating in the presence versus absence of a control system. In contrast, our study recognizes that effective control systems provide economic incentives for cooperation, and so the level of cooperation is likely to be higher when a control system is in place. Thus, our study incorporates the mediating role of control-induced cooperation.

---

[5] Researchers have made similar claims in other contexts related to the more general idea that extrinsic motivation reduces intrinsic motivation (including trust) (Bénabou and Tirole 2003; Bohnet et al. 2001; Frey and Jegen 2001; Pfeffer 1998a, 1998b), a concept formalized in economics as "Motivation Crowding Theory" (Frey 1997). For example, Yamagishi (1998) argues that Japanese subjects are less trusting than American subjects because contract enforcement mechanisms are more prevalent in Japan.

Further, some studies (e.g., Tenbrunsel and Messick 1999) overlook feedback (e.g., performance reports), a common component of formal control systems aimed at reducing relational risk. This limitation is significant, because feedback can potentially mitigate the negative framing effects described above. In the multiperiod collaborative setting, feedback regarding partner actions becomes an additional environmental cue that affects how collaborators evaluate their partners and how they frame the situation in subsequent decision-making periods. Specifically, in viewing their partners' control-induced cooperation, collaborators may favorably update their beliefs about partner trustworthiness. Kramer (1999) notes that perceptions of trustworthiness are largely based on such interactive histories.

In summary, prior studies inhibit the positive effects of control by ignoring the possibility that observed control-induced cooperation can build trust. In contrast to prior work, we predict and show that a strong control system increases cooperation. Further, we allow collaborators to observe this increased cooperation (i.e., cooperation feedback). We expect this observation of control-induced cooperation to lead to increased trust because of a documented tendency of observers to overattribute others' behaviors to dispositional characteristics (Gilbert and Malone 1995; Jones and Harris 1967). This result is so widely observed that it is known as the Fundamental Attribution Error (Ross 1977).

In the current study's context, when collaborators cooperate in the presence of a control system, observers (i.e., other collaborators) are likely to at least partially attribute this control-induced cooperation to the collaborators' inherent trustworthiness. Thus, the control system will have a positive effect on the level of trust among collaborators. Importantly, although collaborators are unlikely to ignore situational factors, the tendency to attribute cooperative behavior to dispositional factors will lead to a strong judgment of trustworthiness (a disposition).

Observed control-induced cooperation provides a basis for the formation of trust between collaborators. This effect is illustrated in Figure 1, Panel B, in two ways. First, the revised relational risk curve lies primarily below the original curve. That is, because the explicit controls are augmented by trust, the probability of opportunistic behavior, and hence, the cost associated with relational risk, decreases for a given level of control.[6] Second, the revised curve is steeper than the original curve. This reflects the notion that the level of cooperation is expected to increase with control strength, and so the positive effects of trust should also increase with control strength.[7] In other words, the marginal benefit of control increases because of the trust that control-induced cooperation engenders.[8] There is a corresponding shift in the total cost curve that results in a higher optimal (i.e., minimum total cost) level of control. The implication is that firms will choose to implement a stronger control system than previous research would seem to suggest. The process by which we predict control to positively affect trust and cooperation is pictured in Figure 2 and summarized by the hypotheses below.

---

[6] This is in contrast to the prevailing view of the effect of control on trust, which would suggest that the revised curve lies above the original curve.

[7] At extreme levels of control system strength, it is theoretically conceivable that the revised curve would not be steeper than the original curve (i.e., the effect of trust on cooperation might not be increasing in the level of control). However, this condition is likely outside the norm of operations for most firms.

[8] We show the revised curve crossing above the original curve at extremely low levels of control to allow for the possibilities (*if the control system is sufficiently weak*) that (1) feedback might reduce trust (because of observed noncooperation), or that (2) the negative framing effects described earlier might outweigh the positive effects we propose.

Hypotheses

As described above, we expect the level of trust to increase in control system strength. Prior studies have generally viewed control as a dichotomous variable (control versus no control). Thus, in developing and testing the hypotheses, we chose two points along the control strength continuum—no control system versus a strong control system. We expect and provide evidence for the necessary condition that the control system will lead to increased levels of cooperation (relative to no control system). We further assume that this control-induced cooperation will be observed by collaborators, providing a mechanism whereby attributions for control-induced behavior can develop. Hypotheses 1a and 1b address the fundamental research question of whether control systems create or destroy trust in collaborative settings. Our theory suggests that when collaborators observe others cooperating, they will at least partially attribute this behavior to the others' inherent trustworthiness. Thus, while two participants may be equally trustworthy, the more cooperative person will be judged as more trustworthy, despite the fact that the cooperation may have been induced by the control system. This results in the following hypothesis.

> **H1a:** Participants will judge others to be more trustworthy when there is a control system in place (that induces cooperation) than when there is no control system.

The mechanism by which this effect occurs is the increased, albeit control-induced, cooperation of participants whose trustworthiness is being assessed. That is, we predict that a strong control system will positively affect perceived trustworthiness precisely because of the control-induced cooperation. Thus, the prediction that control systems increase trust depends critically on the role of observed cooperation as a mediating variable. Accordingly, we make the following prediction:

> **H1b:** The effect of a control system on perceived trustworthiness will be mediated by cooperation.

The trust that develops from observing control-induced cooperation is self-reinforcing in that, consistent with the prior literature, we expect trust to lead to subsequent cooperation. There is a large body of literature in management and psychology claiming that trust promotes cooperation and is, in fact, an alternate control mechanism (Bradach and Eccles 1989; Das and Teng 1998). Most of these studies claim that trusting another individual promotes more cooperative behavior from the *trusted* individual. In this study, we argue further that trust promotes cooperative behavior from the *trusting* individual. Specifically, preferences for fairness, equity, and reciprocity are expected to discourage one from taking advantage of a trusted collaboration partner (Fehr and Falk 2002). Similarly, anticipated feelings of guilt are believed to dissuade one from behaving opportunistically against a trusted collaboration partner.

> **H2:** The level of cooperation will be greater for collaborators with higher perceptions of each other's trustworthiness.

When viewed together, these hypotheses present a self-reinforcing effect of control on cooperation. That is, the marginal benefit of control increases because of the trust that control-induced cooperation engenders. This self-reinforcing control process provides the basis for our fundamental claim that, contrary to prior research findings, control systems can increase trust and cooperation in collaborative settings.

III. EXPERIMENTS

We conducted two experiments. In both experiments we manipulated the presence or absence of a control system. Experiment 1 was conducted in the tradition of psychology-based experimental work, and allowed us to test H1a and H1b, related to the effect of control on cooperation and the effect of cooperation on trust. Experiment 2, which took the approach of experimental economics, allowed us to verify these effects and to test H2, related to the effect of trust on subsequent cooperation. Each experiment involved the interaction of two individuals in a collaborative arrangement. The two experiments are described in detail below, and our operational model is pictured in Figure 2.

Experiment 1

Method

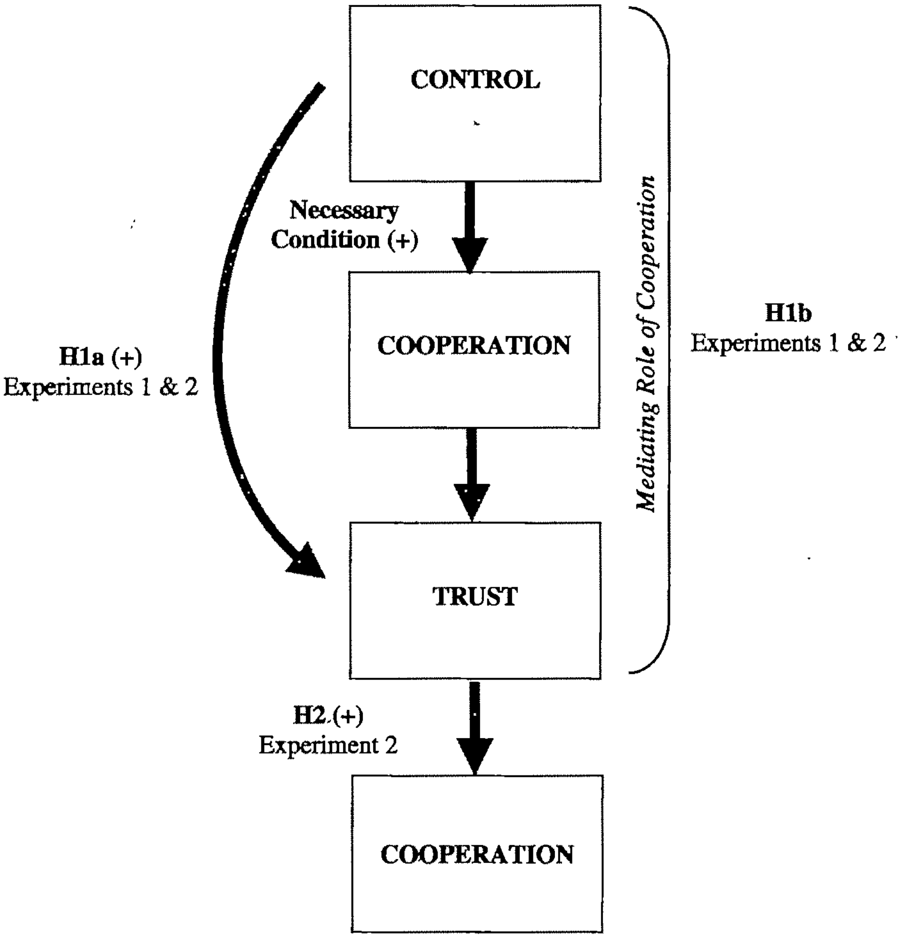
A total of 82 undergraduate (sophomore and junior status) students were recruited for Experiment 1 from an introductory management accounting course at a large state university. Participants received $10 for their participation in this experiment, which involved a two-way between-subjects design. The independent factor is the control system (present or absent). Further, each observation comprises two participants, who fulfill the roles of collaborator and observer. The sessions were conducted in sequence on the same day with the observer sessions following the collaborator sessions. Each session lasted approximately 30 minutes. Participants' roles are discussed below in turn.

Participants in the role of *collaborator* read a scenario, in which they were instructed to assume the role of a research and development (R&D) manager at a large pharmaceutical company. As such, the participants had to decide whether to devote most of the division's limited resources to its individual projects or to a project being conducted jointly with another R&D division. The scenario specified that top management preferred that resources be dedicated to the joint project, and that the participant had previously committed to devote a high level of resources to the joint project. The scenario presented a dilemma, however, in that if the joint project failed, it would be difficult for top management to determine which division was to blame, whereas if the individual projects failed due to lack of resources, the participants would be held personally responsible. While the scenario did not describe an economic game with great specificity, the scenario left little doubt that the principled course of action would be to dedicate one's resources to the joint project because of the previous commitment. Further, investment in the joint project was cooperative, in that the joint project drew on the expertise of both divisions and was beneficial to the firm as a whole (and by simple extension, to the other R&D manager). Thus, social norms would dictate investing in the joint project, and a trustworthy individual would uphold these norms. However, the participants had to weigh these norms against the cost of failing on the individual projects.

Those participants in the *control system* condition were also instructed that a consultant had been hired to make periodic, unannounced visits. If the consultant determined that a division had committed an insufficient level of resources to the joint project, top management would likely penalize that division in the form of reduced budgets in the future, or a lower bonus to the division manager. Participants in the *no control system* condition did not face the potential of consultant visits.

In addition to choosing to dedicate high resources to individual or joint projects, each participant in the role of collaborator wrote an essay that described his/her thought processes and reasoning behind the resource allocation decision. The participants in the observer role then reviewed these collaborator responses to assess the collaborators' trustworthiness. Importantly, collaborators were randomly assigned to control system conditions, such that

FIGURE 2
The Effect of Control on Trust and Cooperation



the inherent trustworthiness of these individuals would not be expected to differ across conditions.

Each *observer* received the full experimental package—scenario and responses—from one collaborator. The observer's task was to read these materials and to make several judgments regarding the trustworthiness of the collaborator. More precisely, observers judged the collaborator's trustworthiness and cooperativeness, and the degree to which the collaborator was a "team player." Recall that we expect (and will verify) that participants in the *control system* condition will cooperate more often than participants in the *no control system* condition. The tendency of individuals to attribute behavior to dispositional characteristics suggests that this control-induced cooperation will lead to positive attributions of the collaborators' trustworthiness. Thus, we predict that observers will judge collaborators in the *control system* condition to be more trustworthy than collaborators in the *no control system* condition (H1a) and that this effect will be mediated by observed cooperation (H1b).

Experiment 1 Results

Table 1 presents descriptive statistics for Experiment 1. We asked a number of manipulation check questions to ensure that participants understood the scenarios and attended to the manipulations. These manipulation checks were answered correctly by substantially all participants, and results are inferentially identical if those participants failing manipulation checks are eliminated.[9] Perhaps most importantly, we asked observer participants to identify (1) the *control system* condition, and (2) the resource allocation choice made by the collaborator participant (i.e., individual or joint projects). For each manipulation check,

TABLE 1
Experiment 1
Descriptive Statistics

Panel A: Collaborator Choices

	Condition		
	No Control System	**Control System**	**Difference U-stat**
Number (%) of participants investing high resources in Joint Project[a]	5 (24%) (n = 21)	15 (75%) (n = 20)	102.5**

Panel B: Observer Judgments

	Condition		
	No Control System	**Control System**	**Difference t-stat**
Mean (Std. Deviation) Ratings of Collaborators[b]			
Trustworthiness	5.62 (1.83)	7.15 (1.73)	2.75**
Team player	4.38 (2.97)	6.75 (2.73)	2.65**
Cooperativeness	5.43 (2.31) (n = 21)	6.70 (2.30) (n = 20)	1.76*

*,** Groups significantly different at $p < 0.05$ and $p < 0.01$, respectively, one-tailed.

[a] Collaborators invested a high level of resources in *either* the joint project *or* individual projects.

[b] Observers rated collaborators' trustworthiness, cooperativeness, and the extent to which the collaborator was a team player on a 1–10 Likert scale.

---

[9] One exception is that ten out of the 41 collaborator participants responded that the case stated they had "previously expressed a willingness to devote a low level of resources to the joint project" when, in fact, the case states that the division head (their assumed role) had expressed a willingness to commit a high level of resources. However, these manipulation check failures appear to have been caused by a misunderstanding of the question. In all but one of these cases, the collaborator's actual choice was to commit high resources to the individual (i.e., not joint) project. Thus, in this manipulation check, these collaborator participants appear to have been reporting the choice they actually made, rather than the prior commitment described in the case scenario (the intended question). This anomaly does not present inferential concerns for several reasons. First, the ten collaborators missing this question were fairly evenly distributed across conditions: four were in the control system condition, and six were in the no control system condition. Second, an examination of the writers' essays provided no evidence of an actual misunderstanding. In neither condition did any writer allude to a belief that he had not made a prior commitment to devote high resources to the joint project. Finally, results are inferentially identical if these participants are eliminated.

all but one observer participant correctly named the corresponding control condition and allocation choice, indicating that observer participants gave reasonable attention to the details of the experiment.

In developing our hypotheses, we take as given that participants will cooperate more in the presence of the control system than when there is no control system. As presented in Table 1, Panel A, our experimental manipulations indeed provide this necessary condition. Recall that in this experiment, a division's choice to devote high resources to the joint project rather than individual projects demonstrates cooperation. In the *no control system* condition, only 24 percent (5/21) of collaborators chose to devote high resources to the joint project. In the *control system* condition, 75 percent (15/20) of collaborators devoted high resources to the joint project. The Mann-Whitney U-test provides statistical evidence that these proportions are significantly different ($U = 102.5$, $p < 0.01$, one-tailed).[10]

Hypotheses 1a and 1b are the primary focus of Experiment 1 and predict that when observing control-induced cooperation, participants will not fully attribute the cooperation to the presence of a control system. Specifically, H1a predicts that those collaborators in the *control system* condition will be judged as more trustworthy than those in the *no control system* condition. Recall that each observer was provided with the complete experimental package (scenario and responses, including the essay) of one collaborator. The observers then answered a series of questions on their perceptions of the collaborators' characteristics. While individual differences in innate trustworthiness may be revealed in the essays, note that participants were randomly assigned to the *control system* versus *no control system* conditions. Thus, the inherent trustworthiness of participants is not expected to vary across conditions so any systematic differences in essays are the result of the control system manipulation.

To test H1a, we rely on the responses of those participants in the role of *observer* (see Table 1, Panel B). The primary dependent variable is (perceived) *Trustworthiness*, the observers' Likert scale (1–10) level of agreement with the statement, "The participant is a trustworthy individual." A t-test finds that *control system* condition collaborators were judged by their observers to be significantly more trustworthy than were the *no control system* condition collaborators (7.15 versus 5.62; $t = 2.75$, $p < 0.01$, one-tailed). This finding is particularly notable because when the observers were asked to recall whether there was a control system in place, only one answered this question incorrectly. Thus, despite the fact that the observers understood the economic incentives facing the collaborators, observers at least partially attributed the collaborators' behavior to their inherent trustworthiness.[11]

Hypothesis 1b predicts that the effect of a control system on perceived trustworthiness will be mediated by cooperation. We conduct a mediation analysis using the three-step method proposed by Baron and Kenny (1986). First, we establish that the independent variable (Control System) has a significant effect on the mediating variable (Cooperation). This is accomplished using the results of the Mann-Whitney U-test reported above (and in

---

[10] Because this scenario is not defined with economic specificity, we cannot compare the actual cooperation levels to any economically rationalized benchmarks. Rather, we can only provide evidence that economic incentives favored cooperation more in the *control system* condition (due to the possibility of economic sanctions if a consultant reported low cooperation) than in the *no control system* condition. As described below, Experiment 2 provides sufficient specification to compare the cooperation levels to what would be expected if an individual were motivated only to maximize wealth.

[11] The H1a result is corroborated with a MANOVA, in which control system (present or absent) is the independent factor. The dependent variables are the Likert scale (1–10) levels of agreement with three statements, "The participant is a trustworthy individual," "The participant is a team player," and "The participant is a cooperative individual." The result is consistent with the univariate test on the "trustworthiness" statement.

Table 1, Panel A), which finds a significant effect of the control system on cooperation ($U = 102.5$, $p < 0.01$, one-tailed). Second, we regress the dependent variable (Trustworthiness) on the independent variable (Control System). This regression is statistically equivalent to the t-test reported above for H1a (and in Table 1, Panel B), and finds that a control system has a significantly positive effect on perceived trustworthiness ($t = 2.75$, $p < 0.01$, one-tailed). Third, we regress the dependent variable (Trustworthiness) on *both* the independent variable (Control System) and the mediator (Cooperation). This multiple regression (not tabulated) finds a significant coefficient on cooperation ($t = 2.05$, $p < 0.05$, one-tailed), and a nonsignificant coefficient on control system ($t = 1.41$, $p = 0.17$, one-tailed).

In sum, once we control for the effect of cooperation on perceived trustworthiness, the direct effect of the control system on perceived trustworthiness is no longer significant. Baron and Kenny (1986) point out the need to compare not only the statistical significance, but also the absolute size of coefficients across regressions. Such a comparison finds that once we control for the effect of cooperation on perceived trustworthiness, the coefficient on the control system becomes smaller (0.88 in the stage 3 regression versus 1.53 in the stage 2 regression). Together, these findings support our H1b prediction that the effect of a control system on perceived trustworthiness is mediated by cooperation. In fact, we find that cooperation *perfectly* mediates the relation, in that the direct effect of the control system on perceived trustworthiness is not significant when the effect of cooperation is controlled.[12]

Experiment 2

Method

While Experiment 1 was designed to explore the effect of control on trust, Experiment 2, which provides a more precisely specified incentive structure, also provides evidence on the effect of trust on subsequent cooperation. Furthermore, since this experiment allowed participants to interact over multiple periods, participants were able to develop richer interactive histories, an important basis for making attributions of trustworthiness (Kramer 1999).

A total of 62 undergraduate (junior and senior status business majors enrolled in finance, marketing, audit and tax classes) and graduate (M.B.A. students enrolled in introductory management accounting classes) students were recruited to participate in Experiment 2. Each participant was assigned to a two-person collaboration, with each collaboration representing an independent observation. While each person was paired with the same collaboration partner throughout the experimental session, these pairings were anonymous. Each person knew only the participant number, but not the identity of the person with whom s/he was paired. Communication among participants was completely restricted, and all decisions were made and recorded via a computer network.

As in Experiment 1, each participant assumed the role of an R&D manager at a large pharmaceutical firm. For each period, s/he decided how much of the division's R&D resources to devote to a joint R&D project. There were only two choices: a high level of resources or a low level of resources. Participants made this decision for 30 separate periods—earning points based on his/her decisions, and those of his/her collaboration partner (i.e., the other division's R&D manager). At the end of the experiment, points were tallied and each participant was paid $1 for every 25 points earned.

---

[12] Commitment Theory (Nesse 2001) alternatively suggests that a control system is a commitment signal if it is self-imposed. Since the control system in our study is imposed exogenously, it is unlikely to provide such a signal. However, the degree to which commitments are upheld is a potential alternate interpretation of observed cooperation, and hence of the mediating variable through which a control system may engender trust.

The experimental parameters are summarized in Exhibit 1. Each participant incurred a cost of 15 points if s/he decided to devote a high level of resources to the joint project.[13] This cost provided an incentive for the participants to devote only a low level of resources to the joint project. However, the two collaboration members shared the income from the joint project, and this income increased with the level of resources dedicated. Each participant's base payoff each period was 50 percent of the joint project's income minus the participant's resource cost. For example, if one participant selected a low level of resources and his/her partner selected a high level of resources, then the participant earned a base payoff of 15 points for that period (0.50*30 − 0 = 15) while the partner earned a base pay of zero (0.50*30 − 15 = 0).

EXHIBIT 1
Experiment 2
Payoff Structure

Panel A: Summary of Payoff Structure

- If either collaborator devoted high resources to the joint project, that collaborator's division was charged 15 points.
- Joint project profits were increasing in the level of resources dedicated to the joint project:
  - If both divisions chose low, joint project profit = 10.
  - If one division chose low and one chose high, joint project profit = 30.
  - If both divisions chose high, joint project profits = 50.
- Joint project profits were shared equally by the two divisions.
- Control System Condition Only—If the consultant visited and determined that a division had contributed high resources, that division would be rewarded 15 points. (Probability of consultant visit = 80% each period.)

Panel B: No Control System Condition—Normal Form Representation Payoff Structure[a]

	Collaborator 2	
Collaborator 1	**low**	**high**
low*	5,5**	15,0
high	0,15	10,10

Panel C: Control System Condition—Normal Form Representation Payoff Structure[b]

	Collaborator 2	
Collaborator 1	**low**	**high**
low*	5,5**	15,12
high	12,15	22,22

* Represents level of resources dedicated to the joint project (low or high).

** Represents the payoff to Collaborator 1 and Collaborator 2, respectively.

[a] Calculated as 50 percent of joint profit minus resource cost (15 points if the division chose high resources).

[b] Calculated as 50 percent of joint profit minus resource cost (15 points if the division chose high resources) plus bonus (80 percent × 15 points if the division chose high resources).

---

[13] This deduction is analogous to the cost (unspecified in amount) of reduced performance on individual projects in Experiment 1.

The design included one independent factor: the presence or absence of a control system. In the *control system* condition, participants were told that an auditor would make random and unannounced visits to each division's site. At the end of these visits the auditor reported to both managers the level of resources each manager contributed to the joint R&D projects. Thus, the control system provided feedback regarding each participant's choices to both participants. The probability of a visit from the auditor was 80 percent in a given period, and if the auditor reported that a manager had selected a high level of resources, this manager received a 15-point bonus.[14] Again, participants were randomly assigned to conditions such that the inherent trustworthiness of these individuals would not be expected to vary across conditions.

Without considering the effects of the control system (the periodic auditor visits and associated bonus), all participants faced a two-party social dilemma (i.e., a prisoner's dilemma), where cooperation is a relevant social norm. Specifically, each participant had individual incentives to dedicate a low level of resources to the joint project, while both participants received a greater payoff if both contributed high resources than if both contributed low resources. However, once the control system is incorporated, the final game is no longer a social dilemma. That is, when there was a control system in place, it was in each player's individual best interest to cooperate (i.e., dedicate a high level of resources). Thus, we again assume that participants in the *control system* condition will cooperate more often than those in the *no control system* condition.

Hypothesis 1a predicts that participants in the *control system* condition will be judged as more trustworthy than those in the *no control system* condition. This prediction depends on the ability of participants to observe each other's control-induced cooperation. Therefore we allow for 20 periods of play before assessing the perceived trustworthiness of each participant's partner. Once participants completed Period 20, the experimenter announced that for the final ten periods the auditor would not visit. Participants were reminded that because the auditor would not be visiting, they would no longer learn their collaboration partners' resource allocation choices. Thus, in the periods after Period 20, the incentives going forward are identical in the *control system* and *no control system* conditions. However, the conditions differ in terms of the interactive histories that have developed.

Before playing the last ten periods, participants judged the likelihood that their collaboration partners would behave cooperatively. Specifically, they answered, "How likely is it that your partner will select a high level of resources these final periods?" on a Likert scale (1–7). Recall that our definition of trust may be conceptualized as the perceived likelihood that another person will cooperate, absent any economic incentives to do so. Because this question was asked after Period 20, when the prospect of an auditor's visit was removed, this question operationalizes our definition of trust.[15] Hypotheses 1a and 1b jointly predict that this perceived likelihood will be greater for those participants in the *control system* condition than for those in the *no control system* condition. Hypothesis 2 predicts that subsequent cooperation will be greater for those collaborations that exhibit greater levels of trust. We test H2 by examining the level of cooperation in the last ten periods.

---

[14] Because Experiment 2 used actual monetary payments, we had to precisely define how this pay would be calculated. In describing the algorithm, we could have used either a bonus or a penalty scheme. We report the results from using a bonus scheme. However, additional experimental data were collected using a penalty scheme. Results from this version of the experiment are inferentially identical to those reported here.

[15] Rather than asking direct questions about the partner's trustworthiness, integrity, etc., we made this question as emotionally neutral as possible so that decisions in the remaining ten periods were not unintentionally affected.

Experiment 2 Results

Table 2 presents descriptive statistics for Experiment 2, and the trends in cooperation are pictured in Figure 3. Because of the computerized nature of the experiment, basic manipulation check questions could be asked after the instructions but before the participants took part in the experiment. These questions took the form of a four-question quiz, and all questions had to be answered correctly before the participant could begin the experiment. Therefore, all participants passed all manipulation checks.

TABLE 2
Experiment 2
Descriptive Statistics

	Condition		
Collaborator Choices	**No Control System**	**Control System**	**Difference t-stat**
Mean (Standard Deviation)			
Cooperation—Periods 1–20[a]	17.67 (6.41)	34.56 (3.85)	8.97**
Average per period (Total/20)	0.88	1.73	
Cooperation—Periods 21–30[b]	7.80 (4.01)	10.88 (5.45)	1.78*
Average per period (Total/10)	0.78	1.09	
Trust[c]	4.23 (1.21) (n = 15)	5.06 (0.91) (n = 16)	2.17*

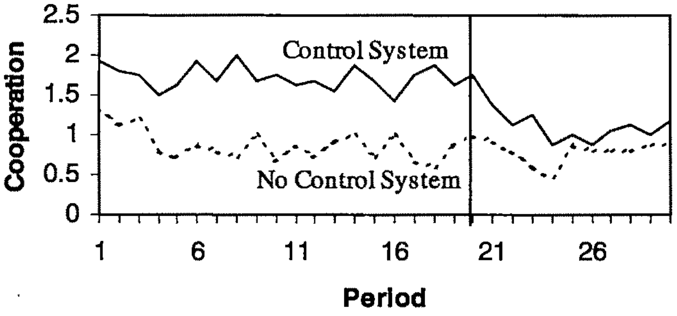
*,** Groups significantly different at $p < 0.05$ and $p < 0.01$, respectively, one-tailed.

[a] Cooperation—Periods 1–20 is the cumulative number of times the two people in a partnership chose to dedicate a high level of resources to the joint project during the first 20 periods (range: 0–40).

[b] Cooperation—Periods 21–30 is the cumulative number of times the two people in a partnership chose to dedicate a high level of resources to the joint project during the last ten periods (range: 0–20).

[c] Trust is the collaborators' assessment of their partners' trustworthiness on a 1–7 Likert scale. Specifically, collaborators answered "How likely is it that your partner will select a high level of resources these final periods?" after being told that no control system would be in place for those final ten periods (periods 21–30).

FIGURE 3
Trends in Experiment 2



The first purpose of Experiment 2 is to corroborate the H1a and H1b results from Experiment 1, using a more precisely specified economic structure. After reporting these results, we proceed to H2, the testing of which is the primary purpose of Experiment 2. We again assume (and test) that cooperation is higher for those participants in the *control system* condition than for those in the *no control system* condition; that is, participants in the *control system* condition will exhibit the control-induced cooperation necessary for trust to develop. The dependent variable for this test is Cooperation—Periods 1–20, the cumulative number of times the two people in a collaboration chose to dedicate high resources to the joint project in the first 20 periods. For a given period, this variable could equal 0, 1, or 2, representing a collaboration in which neither, one, or both partners dedicated high resources to the joint project. Thus, the cumulative variable for all 20 periods ranges from 0 to 40.[16] In Table 2, a t-test confirms that collaborations in the *control system* condition cooperated (chose high resources) more often in the first 20 periods than teams in the *no control system* condition (34.56 versus 17.67; $t = 8.97$, $p < 0.01$ one-tailed), satisfying the necessary condition that the control system induces cooperation.[17]

Hypothesis 1a predicts that other collaborators will positively attribute the documented control-induced cooperation to the collaborators' trustworthiness. Recall that in Experiment 2, after 20 periods of play, participants in the *control system* condition were informed that the auditor would not visit again. Thus, participants in the *control system* and *no control system* condition faced the same "no control system" situation going forward in the final ten periods. After the announcement was made, all participants assessed the likelihood (on a Likert scale) that their collaboration partners would dedicate high resources to the joint project in the final ten periods. This variable, Trust, operationalizes our concept of trust as the perceived likelihood that a partner will cooperate, absent any economic incentives to do so. In Table 2, the responses are averaged for each pair, and a t-test confirms that participants in the *control system* condition expect a higher level of cooperation than those in the *no control system* condition (mean Trust values of 5.06 versus 4.23; $t = 2.17$, $p < 0.05$, one-tailed). This supports H1a, and allows us to conclude that the control system had a positive effect on trust.

Hypothesis 1b predicts that the effect of a control system on trust will be mediated by cooperation. As for Experiment 1, we conduct a mediation analysis using the three-step method proposed by Baron and Kenny (1986). First, we establish that the independent variable (Control System) has a significant effect on the mediating variable (Cooperation). This is accomplished using the t-test reported above (and in Table 2), that documents a significant positive effect of the control system on Cooperation—Periods 1–20 ($t = 8.97$, $p < 0.01$). Second, we regress the dependent variable (Trust) on the independent variable (Control System). This regression is statistically equivalent to the t-test reported above for H1a (and in Table 2), and finds that a control system has a significantly positive effect on

---

[16] The average total collaboration resources in periods 1–20 is significantly greater than the economic prediction of 0 in the *no control system* condition ($t = 10.7$, $p < 0.01$), and significantly less than the economic prediction of 40 in the *control system* condition ($t = 5.7$, $p < 0.01$). Thus, while economic factors play a role, as evidenced by the significant difference in cooperation across conditions, other factors, such as individual preferences and social norms, are also factors in determining the level of cooperation.

[17] Multiple period games such as ours often afford the opportunity for trend analysis (e.g., repeated measures ANOVAs). Such analysis is difficult in our setting because, for each period, the dependent variable is trichotomous, rather than continuous. However, casual observation of changes in cooperation over time reveals no trends, largely because our control system was strong enough to provide near perfect cooperation, even in the earliest periods (e.g., the average level of cooperation in the first five periods was 1.7 out of 2 in the control system condition).

Trust ($t = 2.17$, $p < 0.05$, one-tailed). Third, we regress the dependent variable (Trust) on *both* the independent variable (Control System) and the mediator (Cooperation). This multiple regression (not tabulated) estimates a significant coefficient on Cooperation—Periods 1–20 ($t = 2.26$, $p < 0.01$, one-tailed), and a nonsignificant coefficient on control system ($t = 0.75$, $p = 0.46$, one-tailed). In other words, once we control for the effect of cooperation on trust, the direct effect of the control system on trust is no longer significant. Further, once we control for the effect of cooperation on trust, the coefficient on the control system becomes smaller (0.26 in the stage 3 regression versus 8.45 in the stage 2 regression). Thus, as is true for Experiment 1, we find that cooperation perfectly mediates the effect of the control system on trust (i.e., perceived trustworthiness).

While the H1a, and H1b results from Experiment 1 were corroborated by the Experiment 2 results, the most significant purpose of Experiment 2 was to investigate the effect of trust on subsequent behavior and in the absence of the control system. Hypothesis 2 predicts that the level of cooperation will be greater for collaborators with higher perceptions of each other's trustworthiness. The independent variable for this test is the dependent variable from the test of H1a, the perceived likelihood of cooperation in Periods 21–30 when the possibility of an audit exists in neither condition (i.e., Trust). The dependent variable is Cooperation—Periods 21–30, measured as the cumulative number of times the two people in a partnership chose to dedicate a high level of resources to the joint project during the last ten periods. This variable ranges from 0 to 20, with 20 representing a pair in which both participants chose high resources in each period. A simple regression of Cooperation—Periods 21–30 on Trust (see Table 3) finds a significantly positive effect (coefficient of 2.40; $t = 3.49$, $p < 0.01$, one-tailed). Thus, H2 is supported, and we conclude that, upon the removal of the control system, the level of cooperation is greater for collaborators with higher perceptions of each other's trustworthiness.

The major implication of this finding is that the effect of control on cooperation is reinforced by the trust engendered by control. In periods 21–30, no control system is in place in either condition, and any difference in cooperation must be due to the level of trust generated by the experiences in the earlier periods. From Table 2, we find that removing the control system does lead to an overall decrease in cooperation.[18] Even so, participants in the *control system* condition cooperate (dedicate high resources to the joint project) significantly more in the final ten periods (i.e., following removal of the control system) than participants in the *no control system* condition (10.88 versus 7.80; $t = 1.78$, $p = 0.04$, one-tailed).

A trend analysis presented in Figure 3 highlights that most of the drop-off in cooperation occurred in the first few periods after Period 20, with cooperation levels hitting a low point of 0.9 in Period 24. This rapid decline in cooperation likely reflects a growing appreciation of the change in the incentive structure and its effect on partner cooperation. After this point, cooperation levels seemed to stabilize. Notably, even after the erosion in cooperation when the control system is removed, some residual effect of trust seems to remain. Specifically, cooperation in the final seven periods is higher in the *control system*

---

[18] Considering only the *control system* condition, the average level of cooperation was 1.7 (out of 2.0) in periods 1–20, and 1.1 (out of 2.0) in periods 21–30. The Wilcoxon Signed Ranks test confirms that these proportions are significantly different ($Z = 3.18$, $p < 0.01$). This finding illustrates that participants not only understood that the control system had been removed but altered their behaviors accordingly. To ensure that this difference is due to the removal of the control system, rather than end-of-game effects, we confirm the comparison using only the two, four, six, or eight periods on either side of period 20. In all cases, we find that cooperation decreased significantly after period 20.

TABLE 3
Experiment 2
Regression of Cooperation—Periods 21–30 on Trust

Model: $\text{Cooperation—Periods 21–30}_i = \beta_0 + \beta_1\text{Trust}_i + \epsilon_i$

Dependent Variable: Cooperation—Periods 21–30[a]	**Coefficient (t-statistic)**
Intercept	−1.788 (−0.543)
Trust[b]	2.397** (3.487)
F-statistic	12.152**
Adjusted R^2	27.1%
n (collaboration scores)	31

** Estimate is statistically significant at $p < 0.01$ (one-tailed)

[a] Cooperation—Periods 21–30 is the cumulative number of times the two people in a partnership chose to dedicate a high level of resources to the joint project during the last ten periods (range: 0–20).

[b] Trust is the collaborators' assessment of their partners' trustworthiness on a 1–7 Likert scale. Specifically, collaborators answered "How likely is it that your partner will select a high level of resources these final periods?" after being told that no control system would be in place for those final ten periods (periods 21–30). This Likert scale response is averaged across the two members of each collaborative partnership.

condition than in the *no control system* condition at a marginally statistically significant level ($F = 1.75$, $p = 0.09$, one-tailed). Thus, we conclude that despite the potentially negative effects discussed in the prior literature, control systems can actually enhance trust, and this enhanced trust can increase future cooperation.

IV. CONCLUSION

In summary, this paper investigates the use of control systems to reduce relational risk and increase trust and cooperation in collaborative environments. We provide evidence that in collaborative settings, control induces cooperation, which, in turn, positively affects trust. Specifically, control systems aimed at reducing relational risk promote greater cooperation, which is observed by participating collaborators. This observed cooperative behavior allows collaborators to build trust in one another, and this trust reinforces the positive effects of the control system in eliciting future cooperation. Thus, the effect of control on cooperation is reinforced by the trust that it engenders. This finding is particularly notable, given the large body of literature suggesting just the opposite—that control systems reduce trust.

The findings of this study have several implications for academic research. Prior research suggests that the presence of control systems reduces trust. However, these studies hold cooperation levels constant or do not provide feedback on collaborative cooperation, and, therefore, underestimate the potential benefits of strong control systems. Our study, on the other hand, provides evidence that control systems can, in fact, lead to higher levels of perceived trustworthiness and, ultimately, future cooperation. The mediating role of control-induced cooperation provides the mechanism by which this positive effect occurs. Taken together, the results of this study suggest an increasing marginal benefit of control system strength arising from the trust that observed control-induced cooperation engenders. The implication is that firms will choose to implement a stronger control system than previous research would seem to suggest.

The findings of this research also have implications for the practice of management accounting, specifically, in the design and implementation of control systems in collaborative settings. Note that the trust-building benefits of control exhibited in this paper may not be experienced in collaborative settings without feedback mechanisms—that is, in settings in which control-induced cooperation cannot be observed. This implies that some control mechanisms—namely, monitoring and mutual performance measurement—may be more important in collaborative settings than previously recognized. Moreover, because our theory suggests that the trust benefits of control accrue over *time* (i.e., via observation of control-induced cooperation), the timing of the implementation of control systems may be important. The earlier a strong control system is put into place, the larger are the ultimate trust and cooperation levels achieved. McKnight et al. (1998, 475) suggest that the "robustness" of trust depends, in part, on "structural assurances" related to "the security one feels about a situation because of guarantees, safety nets, or other structures." By providing these structural assurances, strong control systems can provide for high levels of trust early in the collaborative relationship.

These implications for management practice suggest that a closer examination of the effect of control on trust and cooperation over time is warranted. In particular, future research should examine whether the trust engendered by observed control-induced cooperation is self-sustaining. Figure 1 suggests that reductions in control necessarily lead to reductions in trust. However, if trust is self-sustaining, it may be possible to maintain cooperation levels while simultaneously reducing the strength (and, hence, the cost) of the control system. If true, then this would help explain the existence of successful collaborations that appear to be governed by weaker explicit control mechanisms. Furthermore, it would provide even more support for the notion that firms should implement stronger control systems in the earliest stages of a collaboration.

One limitation of the current study is our choice to model only intra-firm collaborations within the experimental design—that is, we focused on teams within a firm. Because our theory applies to collaborations in general, and does not depend on the collaboration existing within the boundaries of a firm, we expect our results to apply to inter-firm collaborations as well. There are, however, differences between intra- and inter-firm collaborations, one of which is the cognitive predisposition of intra-firm collaborators to feel as if they are on the "same team." Indeed, Towry (2003) documents that this sense of "team identity" increases the probability of cooperation. While we know of no theories suggesting that such an effect would negate the findings of the current study, we leave to future research an investigation of the degree to which our findings generalize to inter-firm collaborative settings.

A second potential limitation of the current study is our choice to use a relatively strong control system. The next logical step is to identify the boundary conditions within which a control system will positively affect trust. For example, a weaker control system may provide a feedback mechanism, but fail to provide strong enough incentives for cooperation. In such a case, the feedback may reveal shirking and, as a result, lead to heightened levels of *distrust* among collaborative partners (see Figure 1, Panel B). More generally, a fruitful avenue for future research lies in providing more insight into the role of *information* in strengthening (or weakening) the links between control, trust, and cooperation identified in this paper.

REFERENCES

Allerton, H. 1996. Triple threat. *Training and Development* 50 (5): 16.

Anderson, S. W., and K. L. Sedatole. 2003. Management accounting for the extended enterprise: Performance management for strategic alliances and networked partners. In *Management Accounting in the Digital Economy*, edited by A. Bhimini. London, U.K.: Oxford University Press.

Arya, A., J. Fellingham, and J. Glover. 1997. Teams, repeated tasks, and implicit incentives. *Journal of Accounting and Economics* 23: 7–30.

Balakrishnan, R., N. J. Nagarajan, and K. Sivaramakrishnan. 1998. The effect of property rights and audit information quality on team incentives for inventory. *Management Science* 44 (9): 1193–1204.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Barron, J. M., and K. P. Gjerde. 1997. Peer pressure in an agency relationship. *Journal of Labor Economics* 15 (2): 234–254.

Bénabou, R., and J. Tirole. 2003. Intrinsic and extrinsic motivation. *Review of Economic Studies* 70 (3): 489–520.

Birnberg, J. G., I. H. Frieze, and M. D. Shields. 1977. The role of attribution theory in control systems. *Accounting, Organizations and Society* 2 (3): 189–200.

———. 1998. Control in interfirm co-operative relationships. *Journal of Management Studies* 35 (4): 421–428.

Bohnet, I., S. Huck, and B. S. Frey. 2001. More order with less law: On contract enforcement, trust and crowding. *American Political Science Review* 95: 131–134.

Bradach, J. L., and R. G. Eccles. 1989. Markets versus hierarchies: From ideal types to plural forms. *Annual Review of Sociology* 15: 97–118.

Che, Y-K., and S-W. Yoo. 2001. Optimal incentives for teams. *American Economic Review* 91 (3): 525–543.

Chenhall, R. H., and K. Langfield-Smith. 2003. Performance measurement and reward systems, trust, and strategic change. *Journal of Management Accounting Research* 15: 117–143.

Church, A. H. 1996. From both sides now: The power of teamwork—Fact or fiction? *The Industrial-Organizational Psychologist* 33 (10). Available at: http://www.siop.org/tip/tip.html.

Clement, J. P. III. 1997. Effective collaboration is the key to successful operator/supplier relationships. *Oil and Gas Journal* 95 (41): 47–48.

Das, T. K., and B. Teng. 1996. Risk types and inter-firm alliance structures. *Journal of Management Studies* 33 (6): 827–843.

———, and ———. 1998. Between trust and control: Developing confidence in partner cooperation in alliances. *Academy of Management Review* 23 (3): 491–512.

———, and ———. 1999. Managing risks in strategic alliances. *Academy of Management Executive* 13 (4): 50–62.

———, and ———. 2000. Instabilities of strategic alliances: An internal tensions perspective. *Organization Science: A Journal of the Institute of Management Sciences* 11 (1): 77–101.

———, and ———. 2001. Trust, control, and risk in strategic alliances: An integrated framework. *Organization Studies* 22 (2): 251–283.

Dawes, R. 1980. Social dilemmas. *Annual Review of Psychology* 31: 169–195.

Dumaine, B. 1994. The trouble with teams. *Fortune* 130 (5): 86–90.

Fehr, E., and A. Falk. 2002. Psychological foundations of incentives. *European Economic Review* 46: 687–724.

Frey, B. S. 1997. *Not Just for The Money. An Economic Theory of Personal Motivation.* Cheltenham, U.K.: Edward Elgar.

———, and R. Jegen. 2001. Motivation crowding theory. *Journal of Economic Surveys* 15 (5): 589–611.

Gilbert, D. T., and P. S. Malone. 1995. The correspondence bias. *Psychological Bulletin* 117 (1): 21–38.

Gordon, W., B. Moeller, and D. M. Palmer. 2000. Becoming an alliance partner of choice. *Point of View Report*. Available at: http://www.accenture.com.

Groot, T. L. C. M., and K. A. Merchant. 2000. Control of international joint ventures. *Accounting, Organizations and Society* 25 (6): 579–607.

Haynes, C. M., and S. J. Kachelmeier. 1998. The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature* 17: 97–136.

Ittner, C. D., D. F. Larcker, V. Nagar, and M. V. Rajan. 1999. Supplier selection, monitoring practices, and firm performance. *Journal of Accounting and Public Policy* 18 (3): 253–281.

Jarvenpaa, S. L., K. Knoll, and D. E. Leidner. 1998. Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems* 14 (4): 29–64.

Jones, E. E., and V. A. Harris. 1967. The attribution of attitudes. *Journal of Experimental Social Psychology* 3: 1–24.

———. 1990. *Interpersonal Perception*. Chapters 1–3. New York, NY: MacMillan.

Juran, J. M. 1974. *Quality Control Handbook*. 3rd edition. New York, NY: McGraw-Hill.

Kale, P., M. Singh, and H. Perlmutter. 2000. Learning and protection of proprietary assets in strategic alliances: Building relational capital. *Strategic Management Journal* 21 (3): 217–237.

Kelley, H. H. 1967. Attribution in social psychology. In *Nebraska Symposium on Motivation*, Vol. 15, edited by D. Levine, 192–238. Lincoln, NE: University of Nebraska Press.

Kramer, R. M., and T. R. Tyler. 1996. *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, CA: Sage Publications.

———. 1999. Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology* 50 (1): 569–597.

Kruglanski, A. W. 1970. Attributing trustworthiness in supervisor-subordinate relations. *Journal of Experimental Social Psychology* 6: 214–232.

Landry, J. 1998. The value of trust. *Harvard Business Review* 76 (1): 18–19.

Malhotra, D., and J. K. Murnighan. 2002. The effects of contracts on interpersonal trust. *Administrative Science Quarterly* 47 (3): 534–559.

March, J. G., and C. Heath. 1994. *A Primer on Decision Making—How Decisions Happen*. New York, NY: The Free Press.

McAllister, D. J. 1995. Affect and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal* 38 (1): 24–59.

McKnight, D. H., L. L. Cummings, and N. L. Chervany. 1998. Initial trust formation in new organizational relationships. *Academy of Management Review* 23 (3): 473–490.

Merchant, K., W. A. Van der Stede, and L. Zheng. 2003. Disciplinary constraints on the advancement of knowledge: The case of organizational incentive systems. *Accounting, Organizations and Society* 28 (2–3): 251–286.

Messick, D. M. 1999. Alternative logics for decision making in social settings. *Journal of Economic Behavior & Organization* 39: 11–28.

Meyer, C. 1994. How the right measures help teams excel. *Harvard Business Review* 72 (3): 95–103.

Moser, D. V. 1998. Using an experimental economics approach in behavioral accounting research. *Behavioral Research in Accounting* 10 (Supplement): 94–110.

Nesse, R. M. 2001. *Evolution and the Capacity for Commitment*. New York, NY: Russell Sage Foundation.

Parkhe, A. 1993. Strategic alliance structuring: A game theoretic and transaction cost examination of interfirm cooperation. *Academy of Management Journal* 36 (4): 794–829.

Pfeffer, J. 1998a. Six dangerous myths about pay. *Harvard Business Review* 76 (3): 109–120.

———. 1998b. Seven practices of successful organizations. *California Management Review* 40 (2): 96–124.

Ross, L. 1977. The intuitive psychologist and his shortcomings. In *Advances in Experimental Social Psychology*, Vol. 10, edited by L. Berkowitz, 173–220. San Diego, CA: Academic Press.

Schifrin, M. 2001. Is your company magnetic? *Forbes Best of the Web* (May 21): 16.

Seal, W., J. Cullen, A. Dunlop, T. Berry, and M. Ahmed. 1999. Enacting a European supply chain: A case study on the role of management accounting. *Management Accounting Research* 10 (3): 303–322.

Sewell, G. 1998. The discipline of teams: The control of team-based industrial work through electronic and peer surveillance. *Administrative Science Quarterly* 43 (2): 397–428.

Spagnolo, G. 1999. Social relations and cooperation in organizations. *Journal of Economic Behavior and Organization* 38: 1–25.

Spekman, R. E., J. Spear, and J. Kamauff. 2001. Supply chain competency: Learning as a key component. Working paper, University of Virginia.

Sprinkle, G. B. 2003. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (2–3): 287–318.

Strickland, L. H. 1958. Surveillance and trust. *Journal of Personality* 26: 200–215.

Tenbrunsel, A. T., and D. M. Messick. 1999. Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly* 44: 684–707.

Towry, K. 2003. Control in a teamwork environment: The impact of social ties on the effectiveness of mutual monitoring contracts. *The Accounting Review* 78 (4): 1069–1095.

Waller, W. S. 1995. Decision-making research in managerial accounting: Return to behavioral-economics foundations. In *Judgment and Decision-Making Research in Accounting and Auditing*, edited by R. H. Ashton, and A. H. Ashton. New York, NY: Cambridge University Press.

———. 2001. Behavioral accounting experiments in market and game settings. In *Experimental Business Research*, edited by R. Zwick, and A. Rapoport. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Williamson, O. E. 1993. Calculativeness, trust, and economic organization theory. *Journal of Law and Economics* 36 (April): 453–486.

Yamagishi, T. 1998. The provision of a sanctioning system in the United States and Japan. *Social Psychology Quarterly* 51: 265–71.

Zaheer, A., and N. Venkatraman. 1995. Relational governance as an interorganizational strategy: An empirical test of the role of trust in economic exchange. *Strategic Management Journal* 16 (5): 373–392.

———, B. McEvily, and V. Perrone. 1998. Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science* 9 (2): 141–159.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 501–538

An Evaluation of Accounting-Based Measures of Expected Returns

Peter D. Easton
University of Notre Dame

Steven J. Monahan
INSEAD, Accounting and Control Area

ABSTRACT: We develop an empirical method that allows us to evaluate the reliability of an expected return proxy via its association with realized returns even if realized returns are biased and noisy measures of expected returns. We use our approach to examine seven accounting-based proxies that are imputed from prices and contemporaneous analysts' earnings forecasts. Our results suggest that, for the entire cross-section of firms, these proxies are unreliable. None of them has a positive association with realized returns, even after controlling for the bias and noise in realized returns attributable to contemporaneous information surprises. Moreover, the simplest proxy, which is based on the least reasonable assumptions, contains no more measurement error than the remaining proxies. These results remain even after we attempt to purge the proxies of their measurement error via the use of instrumental variables and grouping. We provide additional evidence, however, that demonstrates that some proxies *are* reliable when the consensus long-term growth forecasts are low and/or when analysts' forecast accuracy is high.

Keywords: *cost of capital; expected rate of return; earnings forecasts; residual income valuation; measurement error.*

I. INTRODUCTION

We develop an empirical approach for evaluating the reliability of estimates of the expected rate of return on equity capital. We use our approach to examine seven accounting-based proxies that are imputed from prices and contemporaneous

---

Workshop participants at the American Accounting Association Annual Meetings, Columbia University, Florida State University, Hong Kong University of Science and Technology, INSEAD, Texas A&M University, The University of Alabama, The University of Arizona, University of Chicago, University of Houston, The University of Iowa, University of Notre Dame, and University of Rochester provided valuable comments on an earlier draft. We thank Ray Ball, Phil Berger, Bruce Johnson, Doron Nissim, Lubos Pastor, Marlene Plumlee, Shiva Sivaramakrishnan, Jim Wahlen, and two anonymous referees for helpful comments and suggestions. Earnings forecast data are from I/B/E/S.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted October 2003
Accepted September 2004

analysts' earnings forecasts.[1] We show that for the entire cross-section of firms, none of these proxies has a positive association with realized returns after controlling for changes in expectations about future cash flows and future discount rates. Moreover, a naïve measure of expected return (the inverse of price to forward earnings) contains no more, and often less, measurement error than the remaining proxies. We provide additional evidence, however, that demonstrates that some proxies *are* reliable when the consensus long-term growth forecasts are low and/or when analysts' forecast accuracy is high.

Similar to the majority of studies in the empirical asset-pricing literature, our inferences about the reliability of a particular expected return proxy are based on its association with realized returns. However, unlike these studies, we assume that realized returns are biased and noisy measures of expected returns. This assumption is motivated by evidence presented in Elton (1999) and Fama and French (2002). These authors demonstrate that "information surprises," which cause realized returns to differ from expected returns, do not cancel out over time or across firms.[2] We show that these information surprises are also correlated with expected returns. Taken together, these observations imply that simple regressions of realized returns on expected return proxies yield spurious inferences because of omitted correlated variables bias.

In light of the above, we adopt an approach that explicitly takes into account the bias and noise in realized returns. Our approach is based on the linear return decomposition developed by Vuolteenaho (2002), who demonstrates that information surprises equal the change in expectations about future cash flows (i.e., cash flow news) less the change in expectations about future discount rates (i.e., return news). Hence, he provides a theoretical foundation for a regression of realized returns on proxies for expected returns, cash flow news, and return news. The coefficients from this regression serve as our initial source of evidence about the reliability of the expected return proxies.

A problem with basing our inferences on the regression coefficients discussed above is that the estimates of these coefficients are affected by errors in variables bias. Because we cannot observe expectations and changes in expectations, all of the regressors (i.e., the expected return proxy, the cash flow news proxy, and the return news proxy) are measured with error. The sign and magnitude of the bias attributable to measurement error is generally unknown when more than one variable in a multivariate regression contains measurement error (e.g., Rao 1973). However, the return decomposition demonstrates that if the components of realized returns are measured without error, then the estimates of the slope coefficients in a regression of realized returns on expected returns, cash flow news, and return news are unambiguously equal to 1. Thus, the bias in the coefficients corresponding to our empirical proxies is well defined. This, in turn, implies that we can modify the econometric method described in Garber and Klepper (1980) and Barth (1991) to develop rankings of measurement error variances of the expected return proxies that serve as our second source of evidence about their reliability.

---

[1] Examples of studies that use accounting-based expected return proxies to evaluate the cross-sectional determinants of expected returns include: Chen et al. (2004), Dhaliwal et al. (2003), Gebhardt et al. (2001), Francis et al. (2003), Gode and Mohanram (2003), Hail (2002), Hail and Leuz (2004), Hribar and Jenkins (2004), and Lee et al. (2003). We illustrate our method by examining variants of several of the expected return proxies used in these studies. However, our method can be used to evaluate any expected return proxy including those based on forecasts of dividends and prices (e.g., Botosan 1997; Botosan and Plumlee 2002; Brav et al. 2004; Francis et al. 2004).

[2] Elton (1999, 1199) states: "The use of average realized returns as a proxy for expected returns relies on a belief that information surprises tend to cancel out over the period of a study and realized returns are therefore an unbiased measure of expected returns. However, I believe there is ample evidence that this belief is misplaced." Fama and French (2002) provide evidence that suggests the abnormally large equity premium observed during the post-war era was attributable to information surprises that took the form of consistent downward revisions in expected future discount rates. We elaborate on these issues in Section II.

Our empirical results suggest that, for the entire cross-section of firms, the accounting based proxies we evaluate are not reliable estimates of the expected rate of return on equity capital. None of the proxies has a positive association with realized returns even after we control for changes in expectations about future cash flows and future discount rates. Moreover, the measurement error variance of a naïve proxy based on the price-to-forward-earnings ratio is never greater, and often lower, than the error variances of the remaining proxies. These results are robust. For example, we evaluate the effectiveness of two commonly used methods for mitigating measurement error: instrumental variables, and grouping. Neither of these approaches lead to improvements in the associations between the expected return proxies and realized returns, nor do they affect the ordering of the measurement error variances (i.e., the measurement error variance of the naïve proxy is no different from the error variances of the remaining proxies).

Further analyses, however, demonstrate that certain proxies are reliable for some subsets of the data. First, we show that the reliability of the proxies is decreasing in the magnitude of consensus long-term earnings growth rate, and that an analog of the measure of expected returns used by Claus and Thomas (2001) is a reliable proxy for a subsample of firms with low consensus long-term growth forecasts. We also demonstrate that when *ex post* analysts' forecasts errors are low, all of the proxies have a positive association with expected returns. Combining these two sets of results with the fact that *ex post* forecast errors are increasing in analysts' long-term growth forecasts leads us to draw two conclusions: (1) the lack of reliability for the general cross-section is partially attributable to low-quality analysts' forecasts, and (2) the consensus long-term earnings growth rate is a useful *ex ante* indicator of reliability (the higher the forecasted growth, rate the lower the reliability).

We contribute to the accounting and finance literatures in three ways. First, we develop an empirical approach for drawing unbiased inferences about an expected return proxy from its association with realized returns even if realized returns are biased and noisy. Second, we provide evidence suggesting that in most circumstances the expected return proxies we evaluate are unreliable. This implies that there is a need for further research on the development of accounting-based measures of expected returns, and that extant evidence based on the proxies we evaluate should be interpreted with caution. Finally, we provide evidence consistent with the notion that the apparent lack of reliability of our expected return proxies is partially attributable to the quality of analysts' earnings forecasts, which suggests that further study of the determinants of analysts' forecast errors is warranted.[3]

Two other studies explicitly aimed at evaluating the reliability of accounting-based measures of expected returns are Botosan and Plumlee (2005) and Guay et al. (2003). Botosan and Plumlee (2005) rank measures of expected returns by comparing coefficients from regressions of expected return proxies on assumed risk factors (e.g., CAPM beta, equity market value, leverage, etc.). While this approach has intuitive appeal, it requires that the researcher make the implicit assumption that the risk factors evaluated are correct and exhaustive, which is unlikely. As discussed in Section II, the return decomposition that serves as the foundation for our tests is based on a tautology: that is, our analyses are based on a fully specified model of the relation between an expected return proxy and *true* expected return.

---

[3] Like other studies (e.g., Abarbanell and Bushee 1997; Francis et al. 2000) that document large valuation errors when analysts' forecasts are used in valuation, we are unable to distinguish between two possible explanations for the measurement error in the expected return proxies: (1) errors in the forecasts, and (2) the restrictive assumptions underlying the implementations of the valuation models used to obtain the estimates of the expected rate of return.

The approach adopted by Guay et al. (2003) is similar to ours as they also evaluate the relation between various expected return proxies and realized returns. There is a crucial difference between our study and theirs, however; we include cash flow news and return news proxies in our regressions, whereas they do not. In particular, the evidence presented in Guay et al. (2003) is based on simple regressions of realized return on the expected return proxy. Hence, as discussed above, the associations they document are likely to be biased measures of the relation between the expected return proxies and true expected return. In Section IV, we provide evidence that bias of this nature exists and is non-trivial.

The remainder of the paper unfolds in the following manner. In the next two sections we describe our empirical method, discuss our proxies of interest, and describe our sample. Our main empirical results are presented in Section IV. The results of our instrumental variables analyses, grouping analyses, and analyses of the relation between the reliability of the expected return proxies and analyst long-term growth forecasts are discussed in Section V. We provide concluding comments in Section VI.

II. EMPIRICAL METHOD

We begin by describing Vuolteenaho's (2002) linear decomposition of realized return into three components: expected return, cash flow news, and return news. Vuolteenaho's (2002) decomposition forms the basis of a linear regression of realized return on the proxies for its components. After discussing this regression we describe how measurement error in the regressors leads to bias in the regression coefficients. Finally, we describe the manner in which we refine the method discussed in Garber and Klepper (1980) and Barth (1991) so that we can isolate the portion of the coefficient bias that is solely attributable to the measurement error in the expected return proxy.

The Components of Realized Returns

Vuolteenaho (2002) demonstrates that firm i's realized, continuously compounded return for year $t+1$, r_{it+1}, can be decomposed into three components: (1) expected return, er_{it+1}, (2) changes in expectations about future cash flows (cash flow news, cn_{it+1}), and (3) changes in expectations about future discount rates (return news, rn_{it+1}). In particular:

$$r_{it+1} \approx er_{it+1} + cn_{it+1} - rn_{it+1}. \tag{1}$$

In Equation (1) expectations underlying er_{it+1} are formed at the end of year t, whereas cn_{it+1} and rn_{it+1} reflect revisions in expectations occurring during year $t+1$.[4] A detailed description of Vuolteenaho's (2002) return decomposition, which is similar to the well-known return decomposition developed by Campbell (1991), is provided in Appendix A. Empirical proxies for expected return, er_{it+1}, cash flow news, cn_{it+1}, and return news, rn_{it+1}, are described in Section III.

Three observations about Equation (1) warrant mentioning. First, Equation (1) is derived from a tautology; hence, our analyses do not rely on implicit or explicit assumptions about investor rationality, the nature of market equilibrium, transactions costs, etc. Second, the linear return decompositions developed by Campbell (1991) and Vuolteenaho (2002) are well accepted. For example, a number of studies in finance use variations of Equation (1) as a means of evaluating the determinants of realized returns (comprehensive literature reviews can be found in Campbell et al. [1997, Chapter 7] and Cochrane [2001, Chapter 20]). Finally, since Equation (1) reflects the effect of changes in expectations about future cash flows and future discount rates on realized returns, it provides a direct means of dealing with Elton's (1999) argument that information surprises cause realized returns to be a biased and noisy measure of expected returns.

---

[4] In Equation (1) the negative sign on return news, rn_{it+1}, reflects the fact that, *ceteris paribus*, increases in future discount rates lead to a decrease in contemporaneous price and, thus, realized return, r_{it+1}, is lower than expected return, er_{it+1}.

The third point is especially pertinent to our study. Bias in realized returns implies that the estimate of the slope coefficient taken from a simple regression of realized returns on expected returns is also biased. If changes in expectations about future cash flows (discount rates) are associated with contemporaneous expected returns, the coefficient on expected returns will be affected by correlated omitted variables bias. This is quite plausible. For example, an explanation for the equity premium puzzle is that during the post-war period the U.S. (and other Western nations) experienced an unprecedented run of "good luck."[5] Hence, the expected future rate of return required by investors as compensation for holding the market portfolio steadily declined (i.e., economy-wide rn_{it+1} was negative), which, per Equation (1), caused the *realized* equity premium to be consistently larger than expected.[6] This, in turn, led to higher than expected realized returns on individual stocks. Moreover, the magnitude of the bias at the individual stock level was arguably increasing in the covariance between a stock's return and the return on the market portfolio. It follows that changes in expectations about future discount rates (i.e., rn_{it+1}) were correlated with both realized and expected returns (i.e., r_{it+1} and er_{it+1}), and the coefficient on er_{it+1} is biased if rn_{it+1} is omitted from the regression.[7]

The Regression Based on Vuolteenaho's (2002) Return Decomposition

We begin our analyses by estimating the following regression for each expected return proxy:

$$r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times \hat{er}_{it+1} + \alpha_{2t+1} \times \hat{cn}_{it+1} + \alpha_{3t+1} \times \hat{rn}_{it+1} + \varepsilon_{it+1}. \tag{2}$$

In Equation (2) $\hat{er}_{it+1}$, $\hat{cn}_{it+1}$, and $\hat{rn}_{it+1}$ represent the expected return proxy, the cash flow news proxy, and the proxy for *negative* return news (i.e., $-1 \times rn_{it+1}$), which we refer to as the return news proxy. The expected return proxies, the cash flow news proxies, and the return news proxies are described in Section III. If these empirical proxies are measured *without* error, α_{1t+1}, α_{2t+1}, and α_{3t+1} are equal to 1 and α_{0t+1} is equal to 0. Hence, one means of evaluating a particular measure of expected returns is to conduct a test of the difference between α_{1t+1} and 1. Unfortunately, these tests do not lead to clear-cut inferences; because we are unable to observe expectations or revisions in expectations, each of the regressors in Equation (2) contains error, which implies the bias in a particular regression coefficient is a complex function of the measurement errors in *all* of the regressors (e.g., Rao 1973). To circumvent this problem we use a refinement of the approach discussed in Garber and

---

[5] See Cochrane (2001, Chapter 21, 460–462) for a discussion of the equity premium puzzle. The discussion under the heading "Luck and a Lower Target" is particularly relevant.

[6] Fama and French (2002) provide specific evidence of this phenomenon.

[7] A similar argument can be made for the inclusion of cn_{it+1} in the regression. In particular, if changes in expectations about future cash flows are correlated with investment opportunities, which, in turn, are correlated with expected returns, cn_{it+1} will be correlated with both realized and expected return. Berk et al. (1999) develop a model in which firms' optimal investment choices are associated with expected returns.

Klepper (1980) and Barth (1991) to isolate the portion of the bias in α_{1t+1} that is solely attributable to the measurement error in $\widehat{er}_{it+1}$.

Measurement Error Analysis

In this subsection we describe the intuition underlying the method we use to isolate the portion of the bias in α_{1t+1} that is solely attributable to measurement error in $\widehat{er}_{it+1}$. Appendix B contains a rigorous description of our econometric approach, which is centered on the following regression:

$$\upsilon_{Cit+1} = \delta_{0t+1} + \delta_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}. \tag{3}$$

In Equation (3) υ_{Cit+1} equals the difference between firm *i*'s observed, realized return for year $t+1$, and the sum of the empirical measures of its components (i.e., $\upsilon_{Cit+1} = r_{it+1} - \widehat{er}_{it+1} - \widehat{cn}_{it+1} - \widehat{rn}_{it+1}$); thus, υ_{Cit+1} equals the combined measurement error in our empirical proxies. Each regressor (i.e., ε^{A}_{1it+1}, ε^{A}_{2it+1}, and ε^{A}_{3it+1}) essentially equals the adjusted error from a regression of one of the proxies on the remaining two (e.g., ε^{A}_{1it+1} is obtained by regressing $\widehat{er}_{it+1}$ on $\widehat{cn}_{it+1}$ and $\widehat{rn}_{it+1}$). Hence, each regression coefficient in Equation (3) measures the relation between the error in a particular proxy and the combined error in all the proxies. The expression for the regression coefficient corresponding to ε^{A}_{1it+1}, which we refer to as the "noise variable," is:[8]

$$\delta_{1t+1} = \sigma^2(\upsilon_{1it+1}) + \{\sigma(\upsilon_{1it+1}, \upsilon_{2it+1}) + \sigma(\upsilon_{1it+1}, \upsilon_{3it+1})\} + \{\sigma(er_{it+1}, \upsilon_{2it+1}) + \sigma(er_{it+1}, \upsilon_{3it+1})\}. \tag{4}$$

In Equation (4) $\sigma^2(\upsilon_{1it+1})$ is the variance of the measurement error in the expected return proxy, $\sigma(\upsilon_{1it+1}, \upsilon_{2it+1})$ denotes the covariance between the measurement error in $\widehat{er}_{it+1}$ and the measurement error in $\widehat{cn}_{it+1}$, $\sigma(\upsilon_{1it+1}, \upsilon_{3it+1})$ is the covariance between the measurement error in $\widehat{er}_{it+1}$ and the measurement error in $\widehat{rn}_{it+1}$, $\sigma(er_{it+1}, \upsilon_{2it+1})$ is the covariance between true expected return and the measurement error in $\widehat{cn}_{it+1}$, and $\sigma(er_{it+1}, \upsilon_{3it+1})$ is the covariance between true expected return and the measurement error in $\widehat{rn}_{it+1}$. If the covariance terms are constant across expected return proxies, then variation across proxies in the noise variable is solely attributable to variation in the measurement error in the expected return proxies (i.e., $\sigma^2(\upsilon_{1it+1})$). This implies that the expected return proxy with the smallest noise variable contains the least measurement error and is the most reliable. Hence, we begin our measurement error analyses by estimating Equation (3) for each of the expected return proxies of interest.

A problem with using the noise variables to rank our expected return proxies is that it is unlikely that the covariance terms shown in Equation (4) are constant across expected return proxies. Because each of our measures of the expected rate of return is based on a unique set of assumptions about dividends, future earnings growth, and terminal profitability, it is likely that the relation between the error in each proxy and the errors in the remaining independent variables is also unique. This implies that inferences based on the relative magnitudes of the different estimates of the noise variable may not provide a meaningful basis for inferring the relative magnitudes of $\sigma^2(\upsilon_{1it+1})$. To circumvent this problem we refine the econometric approach developed by Garber and Klepper (1980) and Barth (1991) and estimate "modified noise variables:"[9]

---

[8] As shown in Appendix B, the expressions for δ_{2t+1} and δ_{3t+1} are similar to the expressions for δ_{1t+1}. Given that our primary interest relates to the measurement error in the expected return proxies, we choose to focus the discussion in the body of the paper on δ_{1t+1}.

$$\begin{aligned}\delta^{M}_{1t+1} = \sigma^2(\upsilon_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} \\ - \{\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})\}.\end{aligned} \tag{5}$$

In Equation (5) $\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})$ equals the covariance between true expected return and the sum of true cash flow news and true return news.

Two observations regarding the modified noise variables are pertinent. First, while there is reason to believe $\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})$ is not equal to zero, it involves only the true values of the constructs, which implies it is constant across proxies. Hence, if differences in $\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})$ are second order, which is a reasonable assumption, differences in the modified noise variables are primarily attributable to differences in the measurement error variances of the proxies (i.e., $\sigma^2(\upsilon_{1it+1})$).[10] Second, δ^{M}_{1t+1} is not a function of the measurement errors in the cash flow news and return news proxies (i.e., υ_{1it+1} and υ_{2it+1}). This implies that rankings based on δ^{M}_{1t+1} are unaffected by risk-related measurement errors in $c\hat{n}_{it+1}$ and $r\hat{n}_{it+1}$.

Summary

To summarize, bias and noise attributable to information surprises imply that simple regressions of realized returns on expected return proxies yield spurious inferences that are attributable to omitted correlated variables. Hence, we include measures of cash flow news and return news in our regressions (i.e., Equation (2)). However, because all of the regressors in Equation (2) contain measurement error, the regression coefficient corresponding to the expected return proxy is not a clear-cut indicator of the reliability of this proxy. To overcome this problem we use a refinement of the method described in Garber and Klepper (1980) and Barth (1991) to estimate the measurement error variances.

III. EMPIRICAL PROXIES AND SAMPLE CONSTRUCTION

Accounting-Based Measures of Expected Return

In this section we provide a brief overview of the seven expected return proxies we evaluate, each of which is imputed from prices and contemporaneous earnings forecasts. Our first proxy is based on the assumption that expected cum-dividend aggregate earnings for the next two years are valuation sufficient. Hence, it essentially equals the inverse of the price-to-forward-earnings ratio. For this reason we refer to it as r_{pe}. The purpose of including r_{pe} in our analyses is to provide a naïve benchmark (based on restrictive assumptions about future earnings growth).[11]

Our next four expected return proxies are each derived from the finite-horizon version of the earnings, earnings growth model developed by Ohlson and Juettner-Nauroth (2003) and described in Easton (2004):

---

[9] The refinements are described in Appendix B.

[10] Two observations support the assumption that differences in $\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})$ across expected return proxies are second order: (1) there is no reason to believe errors in our ability to measure expectations at time t are correlated with revisions in true expectations occurring during time $t+1$, and (2) even if this correlation is non-zero, there is no reason to believe its magnitude differs across proxies.

[11] To be precise, the valuation model underlying r_{pe} relies on the assumption that after year $t+2$ cum-dividend aggregate earnings grow at a rate equal to the cost of capital.

$$
\begin{aligned}
P_{it} &= \frac{eps_{it+1}}{r} + \frac{eps_{it+2} + r \times dps_{it+1} - (1 + r) \times eps_{it+1}}{r \times (r - \Delta agr)} \\
&= \frac{eps_{it+1}}{r} + \frac{agr_{it+1}}{r \times (r - \Delta agr)}. \qquad (6)
\end{aligned}
$$

In Equation (6) P_{it} is price at the end of year t, $eps_{it+\tau}$ is the year t forecast of year $t+\tau$ earnings per share, dps_{it+1} is the year t forecast of dividends paid in year $t+1$, r is the expected rate of return, and Δagr is a growth rate. Following Easton (2004), we refer to the difference between expected year-two cum-dividend accounting earnings (i.e., $eps_{it+2} + r \times dps_{it+1}$) and normal accounting earnings that would be expected given earnings of period one (i.e., $(1 + r) \times eps_{it+1}$) as "abnormal growth in earnings" or agr_{it+1}. Hence, Δagr equals the perpetual rate of change in abnormal growth in earnings beyond the forecast horizon.

The first proxy derived from Equation (6) embeds the assumption that no dividends are paid in year $t+1$ and that Δagr equals 0. As shown in Easton (2004), this proxy is equal to the square root of the inverse of the PEG ratio; hence, we refer to it as r_{peg}.[12] Relaxing the assumption that dps_{it+1} is equal to 0 yields a modified version of the PEG ratio and a proxy we refer to as r_{mpeg}.[13] A criticism of r_{peg} and r_{mpeg} is that the assumption of constant abnormal growth in earnings is too restrictive. Gode and Mohanram (2003) avoid this criticism by assuming Δagr is a cross-sectional constant equal to the difference between the risk-free rate of interest and 3 percent. We refer to their proxy as r_{gm}. Easton (2004), on the other hand, simultaneously estimates r and Δagr for portfolios of stocks allowing for cross-sectional variation in Δagr. We refer to his proxy as $r_{\Delta agr}$.[14]

Our final two measures of expected return are derived from the residual income valuation model. We refer to the first of these two proxies as r_{ct} because it is based on the work of Claus and Thomas (2001). Claus and Thomas (2001) assume that earnings grow at the analysts' consensus long-term growth rate until year $t+5$. They assume earnings after year $t+5$ grow at the rate of inflation, which is set equal the risk-free rate less 3 percent. The second proxy derived from the residual income model is based on the work of Gebhardt et al. (2001); hence, we refer to it as r_{gls}. Gebhardt et al. (2001) use actual earnings forecasts to develop estimates of return on equity for year $t+1$ and $t+2$. They assume that accounting return on equity linearly fades to the historical industry median between years $t+3$ and $t+12$, and remains constant thereafter.[15]

Finally, note that all of our expected return proxies reflect continuous compounding. In particular, the value of $e\hat{r}_{it+1}$ for a particular set of assumptions equals $\ln(1 + r)$, where r is the discount rate implied by the corresponding valuation model. The valuation models underlying each of the expected return proxies and details of their calculation are provided in Table 1.

---

[12] The PEG ratio, which is equal to the PE ratio divided by the short-term earnings growth rate, is a common means of comparing stocks in analysts' reports.

[13] Since I/B/E/S does not provide forecasts of dividends, we assume dps_{t+1} equals dps_t (i.e., dividends per share paid in year t).

[14] This method simultaneously estimates the expected rate of return and the growth rate for a portfolio of firms that have similar PEG ratios. Assigning these estimates of the expected rate of return to each firm in the portfolio introduces measurement error. However, it allows us to avoid having to make *ad hoc* assumptions about Δagr.

[15] While both r_{ct} and r_{gls} are based on the residual income valuation model, as discussed above, each reflects different assumptions about the manner in which *ROE* evolves after year $t+2$. Our empirical results demonstrate that these differences in assumptions lead to significant differences in the statistical properties of the two proxies. For example, as shown in Table 3, the correlation between r_{ct} and r_{gls} is less than 0.50.

TABLE 1
Summary of Empirical Proxies, Data Sources, and Sample Construction

Expected Return Proxies

er^p_{it+1}	Valuation Model	Comments
r_{pe}	$P_{it} = \frac{eps_{it+1} + r \times dps_{it} + eps_{it+2}}{(1+r)^2 - 1}$	
r_{peg}	$P_{it} = \frac{eps_{it+2} - eps_{it+1}}{r^2}$	
r_{mpeg}	$P_{it} = \frac{eps_{it+2} + r \times dps_{it} - eps_{it+1}}{r^2}$	
r_{gm}	$P_{it} = \frac{eps_{it+1}}{r} + \frac{eps_{it+2} + r \times dps_{it} - (1+r) \times eps_{it+1}}{r \times (r - \Delta agr)}$	Δagr is the contemporaneous yield on a ten-year government bond less 3 percent.
$r_{\Delta agr}$	$P_{it} = \frac{eps_{it+1}}{r} + \frac{eps_{it+2} + r \times dps_{it} - (1+r) \times eps_{it+1}}{r \times (r - \Delta agr)}$	r and Δagr are estimated simultaneously via the approach discussed in Easton (2004).
r_{ct}	$P_{it} = bps_{it} + \sum_{\tau=1}^{4} \frac{(ROE_{it+\tau} - r) \times bps_{it+\tau-1}}{(1+r)^{\tau}} + \frac{(ROE_{it+5} - r) \times bps_{it+4} \times (1+\gamma)}{(r-\gamma) \times (1+r)^4}$	$ROE_{it+\tau} = eps_{it+\tau}/bps_{it+\tau-1}$. $eps_{it+\tau} = eps_{it+2} \times (1 + ltg_i)^{\tau-2}\ \forall\ \tau > 2$. ltg_i is the I/B/E/S consensus forecast of the growth rate in earnings per share. $bps_{it+\tau} = bps_{it+\tau-1} + eps_{it+\tau} \times (1 - K)$. For profitable firms $K = \max(0,\min(dps_{it}/eps_{it},1))$. For loss firms $K = \max(0,\min(dps_{it}/(.06 \times bps_{it}),1))$. γ is the contemporaneous yield on a ten-year government bond less 3 percent. We solve for r via an iterative procedure.
r_{gls}	$P_{it} = bps_{it} + \sum_{\tau=1}^{11} \frac{(ROE_{it+\tau} - r) \times bps_{it+\tau-1}}{(1+r)^{\tau}} + \frac{(ROE_{it+12} - r) \times bps_{it+11}}{r \times (1+r)^{11}}$	$ROE_{it+\tau} = eps_{it+\tau}/bps_{it+\tau-1}$ for $\tau = 1,2$. $ROE_{it+\tau} = ROE_{it+\tau-1} - fade\ \forall\ \tau > 2$. $fade = (ROE_{it-2} - HIROE_t)/10$, $HIROE_t$ is the historical industry median ROE for all firm-years in the same industry spanning year $t-4$ through year t with positive earnings and equity book values. We use the industry definitions shown in Fama and French (1997). $bps_{it+\tau} = bps_{it+\tau-1} \times (1 + ROE_{it+\tau} - K)$. For profitable firms $K = \max(0,\min(dps_{it}/eps_{it},1))$. For loss firms $K = \max(0,\min(dps_{it}/(.06 \times bps_{it}), 1))$. We solve for r via an iterative procedure.

(continued on next page)

TABLE 1 (continued)

Cash Flow News Proxy, $c\hat{n}_{it+1}$, and Return News Proxy, $r\hat{n}_{it+1}$

Formula	Comments
$c\hat{n}_{it+1} = (roe_{it} - froe_{it,t}) + (froe_{it+1,t+1} - froe_{it,t+1}) + \frac{\rho}{1 - \rho \times \omega_t} \times (froe_{it+1,it+2} - froe_{it,t+2})$	$roe_{it} = \ln(1+ROE_{it})$, $ROE_{it} = eps_{it}/bps_{it-1}$. $froe_{ij,k} = \ln(1+ROE_{ij,k})$, $ROE_{ij,k}$ denotes the forecast of return on equity for fiscal year k and is based on the consensus I/B/E/S forecast made at the end of year j. The approach we use to estimate ρ is discussed in Appendix A. ω_t is estimated on an annual basis via the pooled, cross-sectional regression shown in Equation (9).
$r\hat{n}_{it+1} = \frac{\rho}{1 - \rho} \times (e\hat{r}_{it+2} - e\hat{r}_{it+1})$	$r\hat{n}_{it+1}$ varies across expected return proxies. The approach we use to estimate ρ is discussed in Appendix A. When conducting our multivariate analyses we define return news as: $r\hat{n}_{it+1} = -\frac{\rho}{1 - \rho} \times (e\hat{r}_{it+2} - e\hat{r}_{it+1})$

We use continuously compounded returns; hence, the value of $e\hat{r}_{it+1}$ for a particular proxy is equal to the natural log of 1 plus the discount rate imputed from the valuation model underlying that proxy. P_{it} is the closing share price for fiscal year t per Compustat (data item 199), dps_{it} is dividends per share for year t per Compustat (data item 26), and bps_{it} (bps_{it-1}) is equity book value at the end of year t ($t-1$) per Compustat (data item 60) divided by common shares outstanding at the end of year t ($t-1$) per Compustat (data item 25). eps_{it} is reported earnings per share for year t per I/B/E/S. eps_{it+1} is the consensus earnings per share forecast for year t+1 per I/B/E/S. When available, we use the actual, consensus forecast for year t+2 as our proxy for eps_{it+2}. When actual forecasts are unavailable we use I/B/E/S forecasts of growth in earnings as the basis for developing our proxy for eps_{it+2}. We eliminate firm-years occurring before 1981 and after 1998, or that do not have a December fiscal year-end. We delete firm-years with missing or non-positive P_{it}, bps_{it}, bps_{it-1}, or common shares outstanding; with negative eps_{it+1}; with $eps_{it+2} < eps_{it+1}$; with $ltg_i < 0$; with missing cum-dividend, continuously compounded stock return in year t+1 per CRSP, r_{it+1}; with a missing cash flow news proxy, $c\hat{n}_{it+1}$; or a missing return news proxy, $r\hat{n}_{it1}$. Finally, firm-years with values of r_{it+1}, $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$ or $r\hat{n}_{it+1}$ in the top or bottom 1/2 percentile of the annual, cross-sectional distribution are considered outliers and deleted. Our final sample consists of 15,680 firm-years.

Cash Flow News and Return News Proxies

In the Vuolteenaho (2002) return decomposition (Equation (1)), cash flow news cn_{it+1} is the component of realized returns corresponding to the change in investors' expectations about future cash flows. As shown in Appendix A, cash flow news is defined as follows:

$$cn_{it+1} = \Delta E_{t+1}\left[\sum_{\tau=1}^{\infty}\rho^{\tau-1} \times roe_{it+\tau}\right]. \tag{7}$$

In Equation (7) $\Delta E_{t+1}[.]$ equals $E_{t+1}[.] - E_t[.]$, ρ is a number slightly less than 1, and *roe* is the natural log of 1 plus the accounting rate of return on equity. This equation implies that a revision in expectations about future profitability leads to a realized return that is larger (in magnitude) than expected return. The effect on realized return of a particular revision in expectations about future profitability is not, however, a cross-sectional constant. Rather, the size of the effect depends on investors beliefs about growth, which are captured by ρ. In particular, as shown in Appendix A, ρ is monotonically increasing in the price-to-dividend ratio, which is generally considered a function of future growth opportunities.[16] Hence, ρ can be viewed as a "capitalization factor."

We use the following formula to estimate our cash flow news proxies:

$$\begin{aligned} \hat{cn}_{it+1} &= (roe_{it} - froe_{it,t}) + (froe_{it+1,t+1} - froe_{it,t+1}) \\ &\quad + \frac{\rho}{1 - \rho \times \omega_t} \times (froe_{it+1,t+2} - froe_{it,t+2}). \end{aligned} \tag{8}$$

In Equation (8) $froe_{ij,k}$ denotes forecasted *roe* for fiscal year k and is based on the consensus forecast of eps_{ik} made in December of year j, ω_t is the expected persistence of *roe* as of time t.[17] The manner in which we estimate ρ, which varies with the price-to-dividend ratio, is discussed in Appendix A. Our cash flow news proxy embeds the assumption that *roe* follows a first order autoregressive process after year $t+1$. This assumption is consistent with evidence presented in Beaver (1970), Freeman et al. (1982), and Sloan (1996).

We estimate ω_t via the following pooled cross-sectional and time-series regression:

$$roe_{i\tau+1} = \omega_{0t} + \omega_t \times roe_{i\tau}. \tag{9}$$

In Equation (9) τ is a number between t and $t-9$ and the sample includes all firm-years in the same Fama and French (1997) industry with requisite data in year $t-9$ through year t (i.e., for each year t we estimate ω_t using firm years between year $t-9$ and year t). Given that accounting methods, competition, and risk vary across industry, our use of an industry specific ω_t reduces the potential for bias attributable to risk-related measurement error in $\hat{cn}_{it+1}$.

---

[16] Consider the dividend growth model $P = \frac{dps}{r-g} \Rightarrow \frac{P}{dps} = \frac{1}{r-g}$, which clearly implies the price-to-dividend ratio is increasing in the growth rate, g.

[17] We include $(roe_{it} - froe_{it,t})$ in Equation (8) to capture the change in expectations that occurs in year $t+1$ upon the announcement of *actual* year t earnings. As discussed below, $froe_{it,t}$ is based on information that becomes publicly available on the third Thursday of the last fiscal month of year t.

In the Vuolteenaho (2002) return decomposition (Equation (1)), return news rn_{it+1} is the component of realized returns corresponding to the change in investors' expectations about future discount rates.[18] As shown in Appendix A, return news is defined as follows:

$$rn_{it+1} = \Delta E_{t+1}\left[\sum_{\tau=2}^{\infty}\rho^{\tau-1} \times r_{it+\tau}\right]. \tag{10}$$

ρ has the same role in the above expression as its role in cn_{it+1} (i.e., it is a capitalization factor that allows the sensitivity of realized return to a given change in the expected discount rate to vary across stocks on the basis of variation in growth). We use the following formula to estimate our return news proxy:

$$r\hat{n}_{it+1} = \frac{\rho}{1-\rho} \times (e\hat{r}_{it+2} - e\hat{r}_{it+1}). \tag{11}$$

That is, $r\hat{n}_{it+1}$ varies across valuation models, and embeds the assumption that changes in the discount rate are permanent (i.e., the discount rate follows a random walk). This assumption is consistent with the fact that each of our expected return proxies is the geometric average of all expected future discount rates.

Our cash flow and return news proxies are a function of analysts' forecasts, which also serve as the basis of our expected return proxies. This is unavoidable as expectations and changes in expectations are inextricably linked. For example, the expected return assigned to a particular stock is a function of the risk of the expected payoffs; hence, cn_{it+1} reflects revisions in the expectations underlying er_{it+1}. Similar logic implies that rn_{it+1} and er_{it+1} are related, which, in turn, implies $r\hat{n}_{it+1}$ varies across valuation models.

Sample Construction

Price at fiscal year-end, equity book value, dividends, and number of shares outstanding are obtained from the 1999 Compustat annual primary, secondary, tertiary, and full coverage research files. Earnings forecasts are derived from the summary 2000 I/B/E/S tape. We determine median forecasts from the available analysts' forecasts on the I/B/E/S file that is released on the third Thursday of December. We delete firms with non-December fiscal year-ends so that the market-implied discount rate and growth rate are estimated at the same point in time for each firm-year observation.[19] For observations in 1995, for example, the December forecasts became available on the 21st day of the month. These data included forecasts for a fiscal year ending ten days later (i.e., December 31, 1995) and either an earnings forecast for each of the fiscal years ending December 31, 1996 and 1997 (i.e.,

---

[18] In the context of a multifactor asset-pricing model, return news is attributable to an unforeseen change in the risk-free rate, unforeseen changes in the factor loadings (i.e., the "betas"), and unforeseen changes in the factor premiums (e.g., the market risk premium). In the instant case, a change in the risk-free rate is irrelevant because it has an equal effect on all equities; hence, it has no effect on the cross-sectional distribution of realized returns. Changes in factor loadings are relevant to the extent they vary across stocks. These changes occur because of changes in the correlation between a particular factor and stock returns, or a change in the volatility of the factor. Changes in factor premiums affect a particular equity by an amount proportionate to the corresponding factor loading for that stock. Hence, even though these changes are economy-wide, they do affect the cross-sectional distribution of realized returns. Changes in factor premiums occur because of changes in investors' appetite for risk, ability to diversify, etc.

[19] While limiting our sample to firms with fiscal-years that end in December reduces the power of our tests, it is unlikely that it leads to bias; hence, our inferences are likely unaffected by this research design choice.

eps_{it+1} and eps_{it+2}) or the forecast for the fiscal year ending December 31, 1996 (i.e., eps_{it+1}) and a forecast of growth in earnings per share for the subsequent years. When available, we use the actual forecasts for the subsequent year (in this example, 1997) as our proxy for eps_{it+2}. When actual forecasts are unavailable we use I/B/E/S forecasts of growth in earnings as the basis for developing our proxy for eps_{it+2}. Realized returns for the year following the earnings forecast date (in the example, January 1, 1996 to December 31, 1996) are obtained from the Center for Research in Security Prices (CRSP) Monthly Return file.

We eliminate firm-year observations (1) prior to 1981 and after 1998 because there are very few observations with complete data in these years, (2) with missing price, dividends per share, or common shares outstanding, (3) with missing or negative book value of equity, (4) with negative eps_{it+1}, (5) with $eps_{it+2} < eps_{it+1}$, (6) with a negative consensus forecast of long-term earnings per share growth, (7) with missing cum-dividend, continuously compounded stock return in year $t+1$, (8) with a missing cash flow news proxy, or (9) with a missing return news proxy. Finally, firm-years with values of r_{it+1}, $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, or $r\hat{n}_{it+1}$ in the top or bottom 1/2 percentile of each yearly, cross-sectional distribution are considered outliers and deleted. Our final sample consists of 15,680 firm-year observations.

IV. MAIN EMPIRICAL RESULTS

In this section we discuss the descriptive statistics, correlations, and the results of our main empirical tests. As discussed above and in Appendix B, the Vuolteenaho (2002) return decomposition pertains to continuously compounded returns; hence, we transform realized returns and each of the measures of expected return by taking the log of 1 plus the proxy of interest. Analyses of untransformed values lead to similar inferences. Descriptive statistics and correlations are based on the definition of $r\hat{n}_{it+1}$ shown in Equation (11). The value of $r\hat{n}_{it+1}$ underlying our multivariate tests is obtained from Equation (11) and is equal to:

$$-\frac{\rho}{1-\rho} \times (e\hat{r}_{it+2} - e\hat{r}_{it+1}).$$

Descriptive Statistics

Descriptive statistics for realized returns, the expected return proxies, and the cash flow news proxy are shown on Panel A of Table 2. The median estimate of r_{pe} is 8.3 percent, which is considerably less than the median realized return of 12.2 percent. Moreover, as shown in the column titled $less\_r_f$, 31 percent of the firm-years in our sample have values of r_{pe} below the contemporaneous value of the continuously compounded risk-free rate. These results suggest that r_{pe} is a downward-biased measure of expected return. However, they do not provide evidence about the cross-sectional association between r_{pe} and true expected returns, which is the pertinent issue given our research question.[20]

---

[20] Given the expected risk premium on a stock is rarely less than zero, a high value of $less\_r_f$ suggests that the *average* of the expected return proxy is a poor measure of *mean* expected returns. However, $less\_r_f$ is not informative about the cross-sectional characteristics of the proxy (i.e., bias does not imply noise). Hence, in the context of our research question a high value of $less\_r_f$ is not *prima facie* evidence that an expected return proxy is unreliable. Conversely, our empirical tests relate to the cross-sectional characteristics of the expected return proxies and do not shed light on whether the proxies are useful measures of mean expected returns. For example, while our evidence suggests that variation in r_{ct} does not capture variation in true expected returns, we *cannot* conclude that the evidence presented in Claus and Thomas (2001) regarding the magnitude of the expected equity premium is unreliable.

TABLE 2
Descriptive Statistics

Panel A: Realized Returns, the Expected Return Proxies, and the Cash Flow News Proxy

	Mean	Std	$less\_r_f$	5th	25th	50th	75th	95th
r_{it+1}	0.096	0.366	0.435	−0.551	−0.093	0.122	0.315	0.631
r_{pe}	0.088	0.034	0.310	0.041	0.065	0.083	0.107	0.152
r_{peg}	0.110	0.032	0.119	0.066	0.090	0.106	0.126	0.169
r_{mpeg}	0.122	0.032	0.026	0.081	0.100	0.116	0.137	0.181
r_{gm}	0.129	0.033	0.012	0.086	0.107	0.124	0.145	0.189
$r_{\Delta agr}$	0.128	0.029	0.005	0.094	0.110	0.122	0.138	0.178
r_{ct}	0.120	0.032	0.020	0.077	0.098	0.115	0.138	0.180
r_{gls}	0.109	0.031	0.105	0.063	0.089	0.107	0.126	0.163
$c\hat{n}_{it+1}$	−0.026	0.384	NA	−0.447	−0.151	−0.037	0.044	0.440

Panel B: The Return News Proxies

	Mean	Std	5th	25th	50th	75th	95th
$r\hat{n}_{it+1}$ (pe)	0.054	1.317	−1.734	−0.298	−0.036	0.265	2.204
$r\hat{n}_{it+1}$ (peg)	−0.027	1.609	−2.390	−0.336	−0.034	0.270	2.352
$r\hat{n}_{it+1}$ (mpeg)	−0.022	1.613	−2.384	−0.359	−0.042	0.297	2.383
$r\hat{n}_{it+1}$ (gm)	−0.047	1.665	−2.461	−0.377	−0.047	0.293	2.343
$r\hat{n}_{it+1}$ (Δagr)	−0.023	1.448	−2.013	−0.291	−0.036	0.233	2.043
$r\hat{n}_{it+1}$ (ct)	−0.076	1.259	−2.104	−0.359	−0.064	0.236	1.909
$r\hat{n}_{it+1}$ (gls)	0.007	0.928	−1.407	−0.274	−0.037	0.233	1.609

r_{it+1} is the realized, continuously compounded return for year $t+1$. r_{pe} is the expected return estimate imputed from the price to forward earnings model. r_{peg} is the expected rate of return implied by the PEG ratio. r_{mpeg} is the expected rate of return implied by the modified PEG ratio. r_{gm} and $r_{\Delta agr}$ are the expected return estimates imputed from Gode and Mohanram's (2003) and Easton's (2004) implementation of the Ohlson and Juettner-Nauroth (2003) model, respectively. r_{ct} and r_{gls} are the expected return estimates imputed from Claus and Thomas's (2001) and Gebhardt et al.'s (2001) implementation of the residual income valuation model, respectively. All the expected return proxies represent continuously compounded returns. $c\hat{n}_{it+1}$ is the cash flow news proxy. $r\hat{n}_{it+1}$ (xxx) is the return news proxy for model xxx. See Table 1 for further details. Mean, Std, 5th, 25th, 50th, 75th, and 95th represent the mean, standard deviation, 5th percentile, 25th percentile, median, 75th percentile, and 95th percentile, respectively. $less\_r_f$ is the proportion of firm-years in which the return variable has a value less than the natural log of 1 plus the contemporaneous risk-free rate of return.

The adjustment provided by taking short-term earnings growth into account (r_{peg}) causes the median estimate of expected returns to increase to 10.6 percent and inclusion of dividends in the expected payoff (r_{mpeg}) causes the median to rise to 11.6 percent. The estimation procedure used by Gode and Mohanram (2003) relies on the assumption that abnormal growth in earnings increases after year $t+1$ and the method used by Easton (2004) leads to estimates of increases in abnormal growth in earnings that are, on average, positive. Hence, the median value of r_{gm} (12.4 percent) and $r_{\Delta agr}$ (12.2 percent) exceeds the median value of r_{mpeg}. The medians of r_{ct} and r_{gls}, which are based on the residual income valuation model, are 11.5 percent and 10.7 percent, respectively.

The median value of the cash flow news proxy ($c\hat{n}_{it+1}$) is −3.7 percent.[21] To the extent that this measure of cash flow news is attributable to analyst optimism (e.g., Richardson et al. 2001) that is ignored by the market, the negative median value of cash flow news represents measurement error in our cash flow news proxies. Moreover, given that our expected return proxies and return news proxies are also derived from analysts' forecasts, the negative median value of cash flow news implies that these constructs are also measured with error.

---

[21] The median (mean) value of ω_t used in the calculation of $c\hat{n}_{it+1}$ is 0.52 (0.55).

Descriptive statistics for our return news proxies are shown in Panel B of Table 2. Since these estimates equal the change in the estimate of expected return over the realized return interval, they differ across the various estimates of expected returns. The median estimates of return news are consistently negative. For example, the median return news implied by the change in r_{peg} is -3.4 percent and the median return news for r_{ct} is -6.4 percent. In light of the fact that prices rose during our sample period, the decline in our expected return proxies suggests a coincident decline in the equity premium.

Correlations

Table 3, Panel A summarizes the correlations among realized returns, the expected return proxies, and the estimates of cash flow news. Pearson product moment (Spearman rank order) correlations are shown above (below) the diagonal. Correlations between our return news proxies and the remaining variables of interest are shown in Panel B. The correlations are the temporal averages of the annual cross-sectional correlations. The t-statistics are the ratio of these averages to their temporal standard errors. We focus our discussion on the Spearman correlations. The Pearson correlations lead to similar inferences.

There is a significant positive correlation between realized return and two of the expected return proxies: r_{ct} (0.073, t-statistic of 2.35) and r_{gls} (0.061, t-statistic of 1.95). None of the correlations between the remaining expected return proxies and realized returns is statistically different from zero at the 0.05 level. Moreover, three of the correlations are negative. These negative correlations imply that the cash flow news and return news components of realized returns reflect more than random measurement error, which only causes attenuation bias and has no affect the sign of the correlation.[22] Rather, these correlations support the arguments we made in Section II. Specifically, true cash flow news and true return news are correlated with our expected return proxies. Hence, in order to avoid drawing spurious inferences attributable to omitted correlated variables bias it is crucial that we control for variation in cash flow news and returns news.

As expected, the correlation between realized returns and the estimate of cash flow news is positive and significant (Spearman correlation of 0.314 with a t-statistic of 16.99). The significant negative correlation between the cash flow news proxy and r_{pe} and the significant negative correlations between the cash flow news proxy and each of the estimates of expected returns derived from the abnormal growth in earnings model (i.e., Equation (6)) provide additional support for our argument that including a cash flow news proxy in our regressions is necessary in order to avoid correlated omitted variables bias. These negative correlations suggest that firms with relatively high discount rates experienced larger

---

[22] The correlation between realized return ($r_{it+1} \approx er_{it+1} + cn_{it+1} - rn_{it+1}$) and an expected return proxy $e\hat{r}_{it+1}$ equals $\frac{\sigma(er_{it+1}, e\hat{r}_{it+1}) + \sigma(cn_{it+1} - rn_{it+1}, e\hat{r}_{it+1})}{\sigma(r_{it+1}) \times \sigma(e\hat{r}_{it+1})}$. If cash flow news and return news are simply random measurement error, then the second term in the numerator equals zero and the sign is equal to the sign of the first term. The sign of the first term is positive unless the covariance between true expected return and the measurement error in the expected return proxy is negative *and* has an absolute value greater than the variance of true expected returns. This is unlikely.

TABLE 3
Correlations among Key Variables

Panel A: Realized Returns, the Expected Return Proxies, and the Cash Flow News Proxy

	r_{it+1}	r_{pe}	r_{peg}	r_{mpeg}	r_{gm}	$r_{\Delta agr}$	r_{ct}	r_{gls}	$c\hat{f}_{it+1}$
r_{it+1}		0.048	−0.032	0.013	−0.006	−0.010	0.050	0.057	0.243
		(1.16)	(−1.40)	(0.54)	(−0.34)	(−0.52)	(1.68)	(2.03)	(13.71)
r_{pe}	0.056		0.384	0.537	0.339	0.325	0.713	0.599	−0.039
	(1.31)		(18.98)	(15.73)	(14.93)	(17.56)	(38.29)	(33.13)	(−2.77)
r_{peg}	−0.028	0.386		0.892	0.880	0.866	0.591	0.415	−0.141
	(−1.13)	(19.48)		(44.39)	(37.33)	(26.11)	(33.64)	(23.05)	(−12.44)
r_{mpeg}	0.023	0.577	0.860		0.956	0.817	0.664	0.394	−0.132
	(0.85)	(18.41)	(38.84)		(60.30)	(20.10)	(39.26)	(19.52)	(−12.94)
r_{gm}	−0.003	0.351	0.843	0.936		0.811	0.567	0.278	−0.134
	(−0.13)	(13.87)	(31.81)	(47.21)		(18.89)	(30.02)	(11.25)	(−13.31)
$r_{\Delta agr}$	0.015	0.360	0.910	0.807	0.795		0.498	0.331	−0.129
	(−0.68)	(19.89)	(101.45)	(38.93)	(32.16)		(20.54)	(14.84)	(−11.89)
r_{ct}	0.073	0.713	0.620	0.724	0.610	0.573		0.428	−0.088
	(2.35)	(40.22)	(37.09)	(59.88)	(28.21)	(38.55)		(14.13)	(−4.55)
r_{gls}	0.061	0.605	0.471	0.451	0.317	0.418	0.470		0.148
	(1.95)	(38.42)	(28.48)	(25.16)	(12.12)	(24.07)	(17.86)		(4.11)
$c\hat{f}_{it+1}$	0.314	−0.030	−0.222	−0.178	−0.191	−0.210	−0.113	0.068	
	(16.99)	(−1.98)	(−20.03)	(−14.93)	(−13.53)	(−19.73)	(−6.16)	(3.43)	

Panel B: Correlations among Return News Proxies, Realized Returns, the Expected Return Proxies, and the Cash Flow News Proxy

	Pearson Product Moment			Spearman Rank Order		
Model	r_{it+1}	$e\hat{r}_{it+1}$	$c\hat{f}_{it+1}$	r_{it+1}	$e\hat{r}_{it+1}$	$c\hat{f}_{it+1}$
$r\hat{n}_{it+1}$ (pe)	−0.339	−0.245	0.055	−0.429	−0.291	0.125
	(−8.97)	(−7.79)	(2.89)	(−14.74)	(−6.49)	(4.81)
$r\hat{n}_{it+1}$ (peg)	−0.174	−0.338	0.031	−0.200	−0.375	0.123
	(−5.36)	(−10.77)	(1.77)	(−7.17)	(−12.00)	(4.49)
$r\hat{n}_{it+1}$ (mpeg)	−0.198	−0.343	0.026	−0.264	−0.387	0.090
	(−6.36)	(−15.79)	(1.48)	(−10.55)	(−18.10)	(3.84)
$r\hat{n}_{it+1}$ (gm)	−0.153	−0.356	0.018	−0.214	−0.402	0.075
	(−5.29)	(−14.91)	(1.10)	(−9.17)	(−17.71)	(3.69)
$r\hat{n}_{it+1}$ (Δagr)	−0.164	−0.414	0.018	−0.174	−0.434	0.097
	(−5.92)	(−11.64)	(1.01)	(−7.38)	(−16.39)	(4.17)
$r\hat{n}_{it+1}$ (ct)	−0.134	−0.286	0.126	−0.223	−0.323	0.251
	(−3.63)	(−11.44)	(8.21)	(−6.58)	(−10.27)	(10.22)
$r\hat{n}_{it+1}$ (*gls*)	−0.444	−0.240	0.024	−0.475	−0.272	−0.001
	(−12.52)	(−9.09)	(1.44)	(−17.03)	(−7.55)	(−0.03)

All correlations reflect averages of the annual correlations. t-statistics (in parentheses) equal the ratio of the correlation to its temporal standard error. Cells above (below) the diagonal of the correlation matrix shown in Panel A correspond to Pearson product moment (Spearman rank order) correlations. The rows in Panel B are the correlations between the measure of return news for a particular model and realized returns, expected returns as measured via the model of interest, and the cash flow news proxy. r_{it+1} is the realized, continuously compounded return for year $t+1$. r_{pe} is the expected return estimate imputed from the price to forward earnings model. r_{peg} is the expected rate of return implied by the PEG ratio. r_{mpeg} is the expected rate of return implied by the modified PEG ratio. r_{gm} and $r_{\Delta agr}$ are the expected return estimates imputed from Gode and Mohanram's (2003) and Easton's (2004) implementation of the Ohlson and Juettner-Nauroth (2003) model, respectively. r_{ct} and r_{gls} are the expected return estimates imputed from Claus and Thomas's (2001) and Gebhardt et al.'s (2001) implementation of the residual income valuation model, respectively. All the expected return proxies represent continuously compounded returns. $c\hat{f}_{it+1}$ is the cash flow news proxy. $r\hat{n}_{it+1}$(xxx) is the return news proxy for model xxx.
See Table 1 for further details.

than average negative information surprises about future cash flows during the time period under study.[23]

The correlation between r_{gls} and the cash flow news proxy is positive and significant, however (Spearman correlation of 0.068 with a t-statistic of 3.43). A rationale for this phenomenon is as follows. There are two essential differences between r_{gls} and the remaining expected return proxies: (1) the residual income valuation model is anchored on year t equity book value, and (2) the terminal value correction used by Gebhardt et al. (2001) is a function of historical industry median return on equity (*ROE*). Thus, compared to the remaining estimates of expected returns, the amount of cross-sectional variation in r_{gls} attributable to variation in the firm-specific book-to-market ratio and variation in industry *ROE* is relatively high. These facts are a likely explanation for the positive correlation between r_{gls} and the estimate of cash flow news: variation in our cash flow news proxy is also a function of variation in equity book value and it is possible that analysts tend to revise their optimistic forecasts of *ROE* toward the industry *ROE*.[24]

Table 3, Panel B summarizes the Pearson and Spearman correlations among our return news proxies and realized return, the cash flow news proxy, and the corresponding estimates of expected return. As expected, the correlation between each of the estimates of return news and realized returns are significantly less than zero (all the t-statistics are less than -3.5). The correlations between the return news proxies and the corresponding proxies for expected returns are all negative and statistically significant. These results are consistent with a decline in the equity premium during our sample period. As discussed in Section II, if the equity premium fell during the sample period, then stocks with high betas (and, thus, high expected returns) experienced larger than average downward revisions in expected future discount rates (i.e., rn_{it+1}).[25] Hence, in order to avoid drawing spurious inferences attributable to omitted variables bias, it is crucial we include a proxy for return news in our regressions.

Multivariate Analyses

Results pertaining to our multivariate tests are shown in Tables 4 and 5. The regression coefficients shown in these tables are the temporal averages of the regression coefficients obtained from annual cross-sectional regressions. The t-statistics are computed via the approach described in Fama and MacBeth (1973).

Summary statistics from the estimation of the regression of realized returns on the empirical measures of its components (expected return, cash flow news, and return news)—Equation (2)—are presented in Table 4 for each of the expected return proxies. The estimates of the coefficient α_1 on the expected return proxies are negative for each of the regressions except the regression based on r_{ct}, and this estimate is not statistically different from zero. Moreover, untabulated results demonstrate that all of the estimates of α_1 are

---

[23] These correlations are also consistent with the notion that analysts' forecasts are too extreme in the sense that high forecasts are too high and low forecasts are too low. For example, if year zero analysts' forecasts are too high (too low), the implied discount rate will be biased upward (downward) and analysts' forecast errors and revisions will be negative (positive).

[24] We test this conjecture by estimating annual regressions of each expected return proxy on the contemporaneous book-to-market ratio and historical industry *ROE*. The average, untabulated R^2s from these regressions are: r_{pe} 0.135, r_{peg} 0.027, r_{mpeg} 0.045, r_{gm} 0.029, $r_{\Delta agr}$ 0.026, r_{ct} 0.05, and r_{gls} 0.225. These results support our conjecture for two reasons: (1) the highest average R^2 is obtained in the regression involving r_{gls}, and (2) the average R^2 is higher when the correlation between the expected return proxy and the cash flow news proxy is higher.

[25] Alternatively, these negative correlations may be attributable to measurement error in our proxies. For example, a decline in analyst optimism may have occurred during our sample period, or extremely high (low) forecasts may be followed by larger than average downward (upward) revisions.

TABLE 4
Multivariate Analyses for Annual Portfolios

Regression: $r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times e\hat{r}_{it+1} + \alpha_{2t+1} \times c\hat{n}_{it+1} + \alpha_{3t+1} \times r\hat{n}_{it+1} + \varepsilon_{it+1}$

$e\hat{r}_{it+1}$	α_0	α_1	α_2	α_3	R^2
r_{pe}	0.16	−0.54	0.29	0.10	0.22
	(5.33)	(−1.46)	(5.28)	(10.36)	
r_{peg}	0.19	−0.67	0.26	0.04	0.12
	(6.26)	(−2.81)	(5.59)	(6.43)	
r_{mpeg}	0.15	−0.36	0.27	0.05	0.13
	(4.02)	(−1.20)	(5.67)	(7.11)	
r_{gm}	0.16	−0.40	0.26	0.04	0.11
	(4.90)	(−1.75)	(5.67)	(5.98)	
$r_{\Delta agr}$	0.20	−0.70	0.26	0.04	0.11
	(6.16)	(−2.70)	(5.70)	(6.41)	
r_{ct}	0.09	0.10	0.30	0.05	0.13
	(1.98)	(0.25)	(5.09)	(4.29)	
r_{gls}	0.24	−1.15	0.29	0.19	0.30
	(7.91)	(−4.21)	(5.52)	(13.52)	

Separate regressions are estimated for each of the 18 annual cross-sections of data. Parameter estimates equal the average of the annual regression coefficients. t-statistics (in parentheses) equal the ratio of the parameter estimates to their temporal standard errors. R^2 is the mean R^2 from the annual regressions. $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, $r\hat{n}_{it+1}$, and r_{it+1} represent the expected return proxy, the cash flow news proxy, the return news proxy and the realized, continuously compounded return for year $t+1$. r_{pe} is the expected return estimate imputed from the price to forward earnings model. r_{peg} is the expected rate of return implied by the PEG ratio. r_{mpeg} is the expected rate of return implied by the modified PEG ratio. r_{gm} and $r_{\Delta agr}$ are the expected return estimates imputed from Gode and Mohanram's (2003) and Easton's (2004) implementation of the Ohlson and Juettner-Nauroth (2003) model, respectively. r_{ct} and r_{gls} are the expected return estimates imputed from Claus and Thomas's (2001) and Gebhardt et al.'s (2001) implementation of the residual income valuation model, respectively. All the expected return proxies represent continuously compounded returns.
See Table 1 for further details.

significantly less than 1 at the 0.01 level. These results suggest there is a considerable amount of measurement error in each of the expected return proxies. On the other hand, as discussed in Section II and Appendix B, the observed bias in α_1 may be attributable to other factors; hence, the need for additional analyses. As expected, the estimates of the coefficients on $c\hat{n}_{it+1}$ and $r\hat{n}_{it+1}$ are significantly positive. For example, for the regression based on r_{gls}, the estimate of the coefficient on $c\hat{n}_{it+1}$ is 0.29 (t-statistic of 5.52) and the estimate of the coefficient on $r\hat{n}_{it+1}$ is 0.19 (t-statistic of 13.52). Nonetheless, untabulated results demonstrate that these estimates are also significantly less than one, which implies that these constructs are also measured with error.

The results of estimating Equation (3) for each of our return proxies are shown in Table 5. Before proceeding with a discussion of the regression coefficients it is interesting to note that Equation (3) explains a considerable portion of the cross-sectional variation in the

TABLE 5
Multivariate Analyses for Annual Portfolios

Regressions: $\upsilon_{Cit+1} = \delta_{0t+1} + \delta_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}$

$\upsilon^{M}_{Cit+1} = \delta_{0t+1} + \delta^{M}_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2it+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}$

	Unmodified Noise Variables (υ_{Cit+1} is the regressand)					Modified Noise Variables (υ^{M}_{Cit+1} is the regressand)				
Model	δ_0	δ_1	δ_2	δ_3	R^2	δ_0	δ^{M}_1	δ_2	δ_3	R^2
pe	0.161	0.0076	0.132	1.214	0.94	0.161	0.0005	0.132	1.214	0.94
	(5.33)	(9.42)	(3.65)	(6.36)		(5.33)	(1.42)	(3.65)	(6.36)	
peg	0.189	0.0152	0.136	2.098	0.96	0.189	0.0012	0.136	2.098	0.96
	(6.26)	(7.70)	(4.01)	(7.89)		(6.26)	(5.92)	(4.01)	(7.89)	
mpeg	0.153	0.0142	0.139	2.093	0.96	0.153	0.0008	0.139	2.093	0.96
	(4.02)	(8.07)	(4.03)	(7.85)		(4.02)	(3.76)	(4.03)	(7.85)	
gm	0.162	0.0161	0.142	2.284	0.96	0.162	0.0010	0.142	2.284	0.96
	(4.90)	(7.99)	(4.08)	(8.32)		(4.90)	(5.31)	(4.08)	(8.32)	
Δagr	0.203	0.0144	0.143	1.712	0.95	0.203	0.0008	0.143	1.712	0.96
	(6.16)	(8.02)	(4.30)	(8.71)		(6.16)	(5.02)	(4.30)	(8.71)	
ct	0.087	0.0070	0.106	1.130	0.93	0.087	0.0003	0.106	1.130	0.93
	(1.98)	(9.48)	(2.91)	(7.30)		(1.98)	(1.25)	(2.91)	(7.30)	
gls	0.239	0.0081	0.150	0.552	0.90	0.239	0.0004	0.150	0.552	0.91
	(7.91)	(9.28)	(4.02)	(7.77)		(7.91)	(1.61)	(4.02)	(7.77)	

Separate regressions are estimated for each of the 18 annual cross-sections of data. Parameter estimates equal the average of the annual regression coefficients. t-statistics (in parentheses) equal the ratio of the parameter estimates to their temporal standard errors. R^2 is the mean R^2 from the annual regressions.

$\upsilon_{Cit+1} = r_{it+1} - e\hat{r}_{it+1} - c\hat{n}_{it+1} - r\hat{n}_{it+1}$, where $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, $r\hat{n}_{it+1}$, and r_{it+1} represent the expected return proxy, the cash flow news proxy, the return news proxy and the realized, continuously compounded return for year $t+1$. All the expected return proxies represent continuously compounded returns. See Table 1 for further details.

$$\upsilon^{M}_{Cit+1} = \upsilon_{Cit+1} - \{\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})\} \times \varepsilon^{A}_{1it+1}$$

$$\varepsilon^{A}_{1it+1} = \frac{-\varepsilon_{1it+1} - \beta_{10t+1}}{\sigma^2(\varepsilon_{1it+1})},\ \varepsilon^{A}_{2it+1} = \frac{-\varepsilon_{2it+1} - \beta_{20t+1}}{\sigma^2(\varepsilon_{2it+1})},\ \text{and}\ \varepsilon^{A}_{3it+1} = \frac{-\varepsilon_{3it+1} - \beta_{30t+1}}{\sigma^2(\varepsilon_{3it+1})}.$$

$\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1})$ $(\sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1}))$ is the covariance between the expected return proxy and the cash flow news (return news) proxy. The β, ε, and $\sigma^2(\varepsilon)$ terms shown above are taken from the following first-stage regressions:

$$e\hat{r}_{it+1} = \beta_{10t+1} + \beta_{11t+1} \times c\hat{n}_{it+1} + \beta_{12t+1} \times r\hat{n}_{it+1} + \varepsilon_{1it+1}$$
$$c\hat{n}_{it+1} = \beta_{20t+1} + \beta_{21t+1} \times e\hat{r}_{it+1} + \beta_{22t+1} \times r\hat{n}_{it+1} + \varepsilon_{2it-1}$$
$$r\hat{n}_{it+1} = \beta_{30t+1} + \beta_{31t+1} \times e\hat{r}_{it+1} + \beta_{32t+1} \times c\hat{n}_{it+1} + \varepsilon_{3it+1}$$

The regression coefficients relating to the measurement error variance of $e\hat{r}_{it+1}$ are:

$$\delta_{1it+1} = \sigma^2(\upsilon_{1it+1}) + \{\sigma(\upsilon_{1it+1}, \upsilon_{2it+1}) + \sigma(\upsilon_{1it+1}, \upsilon_{3it+1})\} + \{\sigma(er_{it+1}, \upsilon_{2it+1}) + \sigma(er_{it+1}, \upsilon_{3it+1})\}$$
$$\delta^{M}_{1t+1} = \sigma^2(\upsilon_{1it+1}) - \{\sigma(er_{it-1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} - \{\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})\}$$

$\sigma^2(\upsilon_{1it+1})$ is the measurement error variance of the expected return proxy, $\sigma(\upsilon_{1it+1}, \upsilon_{2it+1})$ $(\sigma(\upsilon_{1it+1}, \upsilon_{3it+1}))$ is the covariance between the measure error in the expected return proxy and the measurement error in the cash flow news (return news) proxy, $\sigma(er_{it+1}, \upsilon_{2it+1})$ $(\sigma(er_{it+1}, \upsilon_{3it+1}))$ is the covariance between true expected returns and the measurement error in the cash flow news (return news) proxy, $\sigma(er_{it+1}, cn_{it+1})$ $(\sigma(er_{it+1}, rn_{it+1}))$ is the covariance between true expected returns and true cash flow news (return news), and $\sigma(\upsilon_{1it+1}, cn_{it+1})$ $(\sigma(\upsilon_{1it+1}, rn_{it+1}))$ is the covariance between the measurement error in the expected return proxy and true cash flow news (return news).

combined measurement error in our estimates of expected returns (i.e., υ_{CLt+1}).[26] For example, the R^2 based on r_{gls} is 0.90. This implies the return decomposition developed by Vuolteenaho (2002) provides a useful characterization of the components of realized returns.

Turning to the regression coefficients, two implications are immediate. First, all of the expected return proxies contain statistically significant measurement error. For example, the proxy with the highest measurement error variance is r_{gm} (δ_1 of 0.0161, t-statistic of 7.99). Second, while r_{ct} contains the least measurement error (δ_1 of 0.007, t-statistic of 9.48), the simplest proxy, r_{pe}, has only a slightly higher measurement error variance (δ_1 of 0.0076, t-statistic of 9.42). Moreover, untabulated results demonstrate that the measurement error in r_{pe} is not significantly greater than the measurement error in r_{ct} or any of the remaining proxies.[27]

Inferences based on variation in δ_1 are predicated on the assumption that the correlations between the measurement error in our expected return proxies and the measurement error in the remaining proxies are the same for each of the expected return proxies. As discussed in Section II this assumption may not be valid. Thus, we evaluate the modified noise variables. The results of these analyses are summarized on the right-hand side of Table 5.

The expected return proxy with the lowest estimate of δ_1^M is r_{ct} (0.0003, t-statistic of 1.25). However, the coefficient on r_{ct} is only slightly less than the estimate of δ_1^M pertaining to r_{pe} (0.0005, t-statistic of 1.42), and the difference between these two coefficients is statistically insignificant (untabulated t-statistic of 0.99). Moreover, the coefficient on r_{pe} is smaller than each of the coefficients on the remaining proxies except r_{gls}, and the coefficient on r_{gls} is not statistically different from the coefficient on r_{pe}.

Sensitivity Analyses

The results presented in Tables 4 and 5 are striking. None of the expected return proxies we evaluate has a statistically positive association with realized return even though we control for information surprises attributable to changes in expectations about future cash flows and future discount rates. Further, none of the proxies has less measurement error than the simplest proxy, r_{pe}, which is based on a restrictive set of assumptions about future growth and profitability. Taken together these results suggest the proxies we evaluate are unreliable. Given the provocative nature of this conclusion, we discuss the (untabulated) results of three sets of robustness checks.

In our first set of robustness checks we evaluate the sensitivity of our results to assumptions about ρ. The results shown in Tables 4 and 5 are based on the assumption that ρ varies with the price-to-dividend ratio; however, it is possible that ρ varies for other risk-related reasons. To alleviate this concern we re-estimate the regressions underlying Tables 4 and 5 using values of ρ that vary with: (1) size (equity market value), and (2) the debt-to-equity ratio.[28] Results based on these robustness checks are similar to the results shown in Tables 4 and 5. We also obtain similar results to those shown in Tables 4 and 5 when, as per Vuolteenaho (2002), we assume ρ is equal to 0.967 for all firm-year observations.

---

[26] The untabulated means (medians) of υ_{CLt+1} for the expected return proxies are: r_{pe} 0.09 (0.03), r_{peg} −0.01 (0.01), r_{mpeg} −0.02 (−0.02), r_{gm} −0.05 (−0.03), $r_{\Delta agr}$ −0.03 (−0.01), r_{ct} −0.07 (−0.03), and r_{gls} 0.02 (0.01). These amounts suggest that, for several of the expected return proxies, the average (median) combined measurement error is nontrivial. They are not, however, informative about the cross-sectional characteristics of the proxies as bias does not imply noise.

[27] The measurement error in r_{pe} is significantly *lower* than the measurement error in all of the expected return proxies derived from the abnormal growth in earnings model (i.e., Equation (6)). To illustrate how we test the difference between δ_1 for two expected return proxies, consider r_{pe} and r_{peg}. We begin by calculating annual differences between the value of δ_1 pertaining to r_{pe} and the value of δ_1 pertaining to r_{peg}. Next, we calculate the mean and standard error of these annual differences, and form t-statistics by taking the ratio of these two numbers.

[28] We use an approach similar to the approach shown in Appendix B to estimate different values of ρ for size (debt-to-equity) quintiles.

In our second set of robustness checks we evaluate whether our results are attributable to extreme observations. We begin by re-estimating Equation (2) via rank regressions, which yield inferences similar to those based on the results shown on Table 4.[29] Rank regressions cannot be used to estimate δ_1 and $\delta_1{}^M$; hence, to evaluate the sensitivity of our estimates of the measurement error in each proxy we re-estimate δ_1 and $\delta_1{}^M$ after deleting firm-years with values of r_{it+1}, $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, or $r\hat{n}_{it+1}$in the top or bottom fifth percentile of the annual cross-sectional distribution.[30] This modification does not change our inferences.

Finally, we evaluate the robustness of our Fama and MacBeth (1973) t-statistics by calculating t-statistics on precision weighted regression coefficients. Specifically, we divide each annual coefficient by its contemporaneous standard error, and calculate t-statistics by dividing the mean of the precision weighted coefficients by their temporal standard errors. These t-statistics are similar to those shown on Tables 4 and 5.

V. INSTRUMENTAL VARIABLES, GROUPING, AND LONG-TERM GROWTH PARTITIONS

Two common approaches for dealing with measurement error are: (1) instrumental variables, and (2) grouping. In this section we examine the effect these methods have on the reliability of the expected return proxies. Our motivation for this analysis is two-fold. First, if instrumental variables or grouping mitigates the measurement error in the expected return proxies, business practitioners and researchers can use these methods as the basis for developing estimates of expected return. Second, these methods provide an additional means of evaluating the robustness of the results discussed in Section IV. In particular, we demonstrate that for each expected return proxy the estimate of α_1 from Equation (2) is not statistically positive even after we attempt to purge the proxy of its measurement error. In addition, r_{pe} continues to perform as well as less simple proxies.[31]

In this section we also examine how our results vary with the magnitude of analysts' forecasts of long-term growth and with the magnitude of analysts' *ex post* forecast errors. Consistent with La Porta (1996) and Frankel and Lee (1998) we demonstrate a positive association between analysts' forecasts of the long-term earnings growth rate, ltg_i, and errors in analysts' forecasts of earnings. Next, by combining evidence about the variation in α_1 and $\delta_1{}^M$ across partitions of the data formed on the basis of ltg_i with evidence about the variation in α_1 and $\delta_1{}^M$ across partitions of the data formed on the basis of *ex post* analysts' forecast errors, we demonstrate that: (1) the results documented in Section IV are attributable to errors in the earnings forecasts underlying our estimates of expected return, and (2) ltg_i is a useful *ex ante* indicator of forecast quality. This evidence is useful for three reasons. First, it sheds light on potential future research opportunities (e.g., development of better earnings forecast models or statistical approaches to purge the error from analysts' forecasts).[32] Second, this evidence provides researchers with guidance regarding the reliability of the expected return proxies in different empirical settings. For instance, researchers should be particularly careful when interpreting evidence about the cross-sectional determinants of the expected return proxies when this evidence is based on a sample that primarily consists of high ltg_i firms. Finally, we show that for observations in the bottom third of the distribution of ltg_i, r_{ct} contains very little measurement error, which suggests this proxy is a reliable measure of expected returns for a nontrivial subset of our sample.[33]

---

[29] Ranks are determined separately for each annual cross-section and the t-statistics pertaining to the rank regressions are estimated via the Fama and MacBeth (1973) procedure.

[30] Valid comparisons of δ_1 ($\delta_1{}^M$) across expected return proxies can only be drawn if the same set of observations underlies each estimate of δ_1 ($\delta_1{}^M$). Hence, for the purposes of this robustness check we discard observations with values of $e\hat{r}_{it+1}$ or $r\hat{n}_{it+1}$ in the top or bottom fifth percentile for *any* of the expected return proxies when estimating δ_1 (δ_1^M) for a *particular* expected return proxy. This leads to a considerably smaller sample size of 8,748 firm-year observations.

[31] We also evaluate a proxy, r_{avg}, which equals the average of r_{pe}, r_{peg}, r_{mpeg}, r_{gm}, $r_{\Delta agr}$, r_{ct}, and r_{gls}. Untabulated results demonstrate that the estimate of α_1 corresponding to r_{avg} is negative, and r_{avg} does not have a lower measurement error variance than r_{pe}.

Instrumental Variables

We implement the instrumental variables procedure in the following manner. First, each of the expected return proxies is regressed on instruments that are for the purposes of these analyses assumed to be correlated with true expected return but uncorrelated with the measurement error. Next, the fitted values from the instrumental variables regression are evaluated in the same manner as the underlying expected return proxies.

We select the following instruments: CAPM beta, market capitalization, the ratio of equity book value to equity market value, the standard deviation of past returns, and industry type. Before discussing the empirical results pertaining to the instrumental variables analysis, we briefly motivate our choice of instruments.

Well-known theoretical results developed by Sharpe (1964), Lintner (1965), and Mossin (1966) suggest that beta is an appropriate measure of systematic risk. Our firm-year estimates of beta, $BETA_{it}$, are derived from regressions of monthly returns on contemporaneous excess returns (i.e., market return less the yield on a one-month treasury bill) where the data are drawn from the 60-month period prior to the earnings forecast date. Malkiel and Xu (1997) focuses on total risk as measured by the variance of returns. We use the standard deviation of the daily returns occurring during the year leading up to the forecast date, SD_{it}, as our measure of total risk. A variety of studies (e.g., Banz 1981; Fama and French 1992) document a negative association between size and realized returns suggesting that market capitalization as of the earnings forecast date, $SIZE_{it}$, is a potential proxy for risk. Consistent with Fama and French (1992) and Berk et al. (1999), we use the ratio of the book value of common equity to the market value of common equity at the earnings forecast date, BM_{it}, as another measure of risk. Finally, a number of papers (e.g., Fama and French 1997; Gebhardt et al. 2001) show that the equity premium differs across industries. Similar to Gebhardt et al. (2001) we use the average of the firm-specific estimates of expected return, IND_{it}, as our proxy for industry type (industry classifications are the same as those in Fama and French [1997]).

The results of estimating regressions of each of the expected return proxies on the four risk proxies ($BETA_{it}$, SD_{it}, $SIZE_{it}$, and BM_{it}) are shown on the left-hand side of Table 6. For all of the expected return proxies other than r_{pe} and r_{ct} the estimates of the coefficients on $BETA_{it}$ and SD_{it}, are positive (the coefficient on $BETA_{it}$ is not statistically significant in the r_{mpeg} and r_{gls} regressions). Contrary to expectations, the estimates of the coefficient on $SIZE_{it}$ are not significantly less than zero (at the 0.05 level) for all proxies other than r_{pe} (-0.000 with a t-statistic of -3.77). A possible explanation for these results is that the I/B/E/S sample on which our analyses are based tends to include larger firms. Finally, all of the proxies except r_{ct} have a positive and statistically significant association with BM_{it}.

---

[32] For example, Guay et al. (2003) provide evidence consistent with the notion that measurement error in accounting based estimates of the expected rate of return is partially attributable to sluggishness in analysts updating of earnings forecasts.

[33] Inferences based on the results obtained from the use of instrumental variables, grouping, ltg_i partitions, and partitions based on analysts' forecast errors are not sensitive to the robustness checks discussed in Section IV.

TABLE 6
Instrumental Variables (IV) Regressions

Regressions:

$$er^P_{it+1} = \theta_{0t} + \theta_{1t} \times BETA_{it} + \theta_{2t} \times SD_{it} + \theta_{3t} \times SIZE_{it} + \theta_{4t} \times BM_{it} + \psi_{it} \quad \text{(IV Regression One)}$$

$$er^P_{it+1} = \theta_{0t} + \theta_{1t} \times BETA_{it} + \theta_{2t} \times SD_{it} + \theta_{3t} \times SIZE_{it} + \theta_{4t} \times BM_{it} + \theta_{5t} \times IND_{it} + \psi_{it} \quad \text{(IV Regression Two)}$$

	Regression One (excluding *IND*)						Regresson Two (including *IND*)						
er^P_{it+1}	θ_{0t}	θ_{1t}	θ_{2t}	θ_{3t}	θ_{4t}	R^2	θ_{0t}	θ_{1t}	θ_{2t}	θ_{3t}	θ_{4t}	θ_{5t}	R^2
r_{pe}	0.096	−0.005	−0.419	−0.000	0.023	0.16	0.002	−0.001	−0.051	0.000	0.016	0.905	0.42
	(12.27)	(−2.15)	(−3.55)	(−3.77)	(8.53)		(0.77)	(−0.96)	(−0.87)	(3.53)	(6.89)	(75.46)	
r_{peg}	0.083	0.010	0.716	−0.000	0.008	0.13	−0.003	0.006	0.531	−0.000	0.007	0.843	0.27
	(28.94)	(5.20)	(12.37)	(−0.34)	(4.62)		(−2.05)	(5.63)	(13.03)	(−2.39)	(4.48)	(44.43)	
r_{mpeg}	0.112	0.002	0.175	0.000	0.013	0.07	−0.002	0.002	0.211	0.000	0.010	0.916	0.21
	(21.24)	(1.15)	(2.15)	(0.55)	(6.81)		(−0.76)	(2.26)	(4.07)	(0.90)	(5.83)	(51.27)	
r_{gm}	0.117	0.003	0.304	0.000	0.010	0.06	−0.001	0.002	0.250	0.000	0.008	0.914	0.17
	(24.17)	(3.33)	(4.57)	(1.45)	(5.34)		(−0.54)	(3.60)	(5.14)	(0.33)	(5.05)	(48.45)	
$r_{\Delta agr}$	0.111	0.006	0.452	−0.000	0.008	0.09	0.003	0.004	0.349	−0.000	0.007	0.862	0.19
	(32.41)	(4.66)	(7.95)	(−0.69)	(5.09)		(1.44)	(5.25)	(8.37)	(−2.00)	(4.86)	(44.50)	
r_{ct}	0.126	−0.001	0.058	−0.000	−0.004	0.08	0.002	0.001	0.149	0.000	−0.006	0.969	0.26
	(15.55)	(−0.51)	(0.56)	(−0.93)	(−1.96)		(0.53)	(0.52)	(2.15)	(2.11)	(−2.89)	(46.54)	
r_{gls}	0.096	0.002	0.190	0.000	0.021	0.17	−0.007	0.002	0.128	0.000	0.019	0.924	0.40
	(14.43)	(1.28)	(1.92)	(1.38)	(9.95)		(−3.20)	(3.25)	(1.73)	(2.16)	(9.23)	(74.14)	

Separate regressions are estimated for each of 18 annual cross-sections of data. Parameter estimates equal the average of the annual regression coefficients. t-statistics (in parentheses) are the ratio of the parameter estimates to their temporal standard errors. R^2 is the average of the annual R^2 statistics. *BETA* is the capital asset pricing model beta estimated over the 60 months prior to the end of year *t*, *SD* is the standard deviation of daily returns in year *t*, *SIZE* is the market capitalization at the end of year *t*, *BM* is the ratio of the book value of common equity to the market value of common equity at the end of year *t*, and *IND* is the mean at the end of year *t* of the expected return proxies across firms in the same industry group. r_{pe} is the expected return estimate imputed from the price to forward earnings model. r_{peg} is the expected rate of return implied by the PEG ratio. r_{mpeg} is the expected rate of return implied by the modified PEG ratio. r_{gm} and $r_{\Delta agr}$ are the expected return estimates imputed from Gode and Mohanram's (2003) and Easton's (2004) implementation of the Ohlson and Juettner-Nauroth (2003) model, respectively. r_{ct} and r_{gls} are the expected return estimates imputed from Claus and Thomas's (2001) and Gebhardt et al.'s (2001) implementation of the residual income valuation model, respectively. All the expected return proxies represent continuously compounded returns.
See Table 1 for further details.

It is evident from the results in the right-hand side of Table 6 that industry type captures much of the variation in the expected returns proxies. For example, the estimate of the coefficient on IND_{it} for the regression where the dependent variable is r_{ct} is 0.969 with a t-statistic of 46.54. Nonetheless, the coefficients on $BETA_{it}$, SD_{it}, $SIZE_{it}$, and BM_{it} are essentially unchanged with the exception that r_{pe} and r_{ct} are positively associated with $SIZE_{it}$ after IND_{it} is included in the regression.

As shown on Table 7, the fitted values taken from the instrumental variables regressions have lower measurement error variances (i.e., δ_1^M) than the raw proxies. On the other hand, only r_{ct} has a fitted value with a lower measurement error variance than the fitted value of r_{pe}, and untabulated results demonstrate that this difference is only statistically significant when the fitted values are taken from regressions that include IND_{it} (t-statistic of 3.55). Moreover, none of the estimates of α_1 shown on Table 7 are statistically greater than zero. Hence, even after attempting to purge the proxies of their measurement error via instrumental variables, we continue to find little evidence supporting an argument that the proxies we evaluate are reliable measures of expected returns.

Several studies (e.g., Botosan and Plumlee 2005; Gebhardt et al. 2001; Gode and Mohanram 2003) use coefficients from regressions of proxies for expected returns on various risk factors (or correlations between proxies for expected returns and risk factors) as means of ranking expected returns proxies. It is evident from the coefficient estimates on the right-hand side of Table 6 that the expected return proxy that is *best* according to the criteria adopted by these authors is r_{peg}. Specifically, in the r_{peg} regression, the t-statistic for the estimated coefficients on $BETA_{it}$ and SD_{it} (5.63 and 13.03, respectively) are higher than the corresponding t-statistics in the other regressions, the estimate of the coefficient on $SIZE_{it}$ is negative (t-statistic of -2.39), and the coefficients on BM_{it} and IND_{it} are highly significant. Nonetheless, the evidence in Table 7 demonstrates that r_{peg} is one of the *worst* proxies as it has the highest δ_1^M and one of the lowest estimates of α_1. In other words, ranking expected returns on the basis of correlations with potential risk factors may lead to erroneous conclusions.

Grouping

The instrumental variables procedure combines variables that are, under the maintained hypothesis, expected to be correlated with expected returns. Grouping, on the other hand, focuses on each variable separately. We group observations into annual portfolios of 10 based on variables designed to minimize the within portfolio variation in the construct of interest (e.g., beta) and maximize the variation across portfolios.[34] All analyses that are conducted at the firm-specific level are repeated using portfolio averages (i.e., average realized return, average expected return, average cash flow news, and average return news).

The results of our portfolio level analyses shown in Table 8 lead to inferences that are similar to those based on the results shown in Tables 4, 5, and 7. In particular, none of the estimates of α_1 taken from Equation (2) is statistically greater than zero. Moreover, the rankings of δ_1^M pertaining to the portfolio averages of the expected return proxies are, by and large, similar to the rankings of δ_1^M pertaining to the firm-specific estimates. There is one exception, however: the simplest expected return proxy, r_{pe}, does not have the lowest

---

[34] Analyses based on portfolios of size 20 lead to similar inferences. Nonetheless, it is possible that the use of even larger portfolios leads to a significant reduction in measurement error. Unfortunately, we do not have sufficient degrees of freedom to reliably evaluate this assertion; hence, we cannot make definitive statements about the reliability of the proxies at the portfolio level. Furthermore, we cannot assess the accuracy of estimates of the expected rate of return that are specifically designed for large portfolios of observations (such as, Easton et al. 2002; O'Hanlon and Steele 2000).

TABLE 7
Multivariate Analyses for Annual Portfolios Using Fitted Estimates of Expected Returns from Instrumental Variables (IV) Regressions

Regressions:

$$r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times e\hat{r}_{it+1} + \alpha_{2t+1} \times c\hat{n}_{it+1} + \alpha_{3t+1} \times r\hat{n}_{it+1} + \varepsilon_{it+1}$$

$$v^{M}_{Cit+1} = \delta_{0t+1} + \delta^{M}_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}$$

	α_1		δ_1^M	
Model	**IV Regression One**	**IV Regression Two**	**IV Regression One**	**IV Regression Two**
pe	−1.40 (−1.63)	−0.69 (−1.16)	0.0001 (0.79)	0.0003 (1.11)
peg	−3.49 (−3.23)	−1.77 (−2.58)	0.0005 (3.59)	0.0006 (3.87)
mpeg	−2.89 (−1.54)	−1.38 (−1.53)	0.0002 (2.28)	0.0003 (2.21)
gm	−7.36 (−3.19)	−1.79 (−2.09)	0.0002 (3.02)	0.0003 (2.98)
Δagr	−4.69 (−2.09)	−2.76 (−2.62)	0.0003 (3.20)	0.0004 (3.34)
ct	0.00 (0.00)	0.84 (1.13)	−0.0001 (−1.28)	−0.0001 (−0.40)
gls	−2.31 (−3.22)	−1.19 (−2.52)	0.0003 (3.87)	0.0005 (2.83)

Separate regressions are estimated for each of the 18 annual cross-sections of data. Parameter estimates equal the average of the annual regression coefficients. t-statistics (in parentheses) equal the ratio of the parameter estimates to their temporal standard errors. R^2 is the mean R^2 from the annual regressions.

$v_{Cit+1} = r_{it+1} - e\hat{r}_{it+1} - c\hat{n}_{it+1} - r\hat{n}_{it+1}$, where $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, $r\hat{n}_{it+1}$, and r_{it+1} represent the predicted expected return proxy taken from the following instrumental variables regressions:

$$e\hat{r}_{it+1} = \theta_{0t} + \theta_{1t} \times BETA_{it} + \theta_{2t} \times SD_{it} + \theta_{3t} \times SIZE_{it} + \theta_{4t} \times BM_{it} + \psi_{it} \quad \text{(IV Regression One)}$$

$$e\hat{r}_{it+1} = \theta_{0t} + \theta_{1t} \times BETA_{it} + \theta_{2t} \times SD_{it} + \theta_{3t} \times SIZE_{it} + \theta_{4t} \times BM_{it} + \theta_{5t} \times IND_{it} + \psi_{it}$$

(IV Regression Two)

In the above regressions *BETA* is the capital asset pricing model beta estimated over the 60 months prior to the end of year *t*, *SD* is the standard deviation of daily returns in year *t*, *SIZE* is the market capitalization at the end of year *t*, *BM* is the ratio of the book value of common equity to the market value of common equity at the end of year *t*, and *IND* is the mean at the end of year *t* of the expected return proxies across firms in the same industry group.

The cash flow news proxy, the return news proxy, and the realized, continuously compounded return for year $t+1$. All the expected return proxies represent continuously compounded returns.

See Table 1 for further details.

$$v^{M}_{Cit+1} = v_{Cit+1} - \{\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})\} \times \varepsilon^{A}_{1it+1}$$

$$\varepsilon^{A}_{1it+1} = \frac{-\varepsilon_{1it+1} - \beta_{10t+1}}{\sigma^2(\varepsilon_{1it+1})},\ \varepsilon^{A}_{2it+1} = \frac{-\varepsilon_{2it+1} - \beta_{20t+1}}{\sigma^2(\varepsilon_{2it+1})},\ \text{and}\ \varepsilon^{A}_{3it+1} = \frac{-\varepsilon_{3it+1} - \beta_{30t+1}}{\sigma^2(\varepsilon_{3it+1})}$$

(*continued on next page*)

TABLE 7 (continued)

$\sigma(\hat{er}_{it+1}, \hat{cn}_{it+1})$ $(\sigma(\hat{er}_{it+1}, \hat{rn}_{it+1}))$ is the covariance between the expected return proxy and the cash flow news (return news) proxy. The β, ε, and $\sigma^2(\varepsilon)$ terms shown above are taken from the following first-stage regressions:

$$\hat{er}_{it+1} = \beta_{10t+1} + \beta_{11t+1} \times \hat{cn}_{it+1} + \beta_{12t+1} \times \hat{rn}_{it+1} + \varepsilon_{1it+1}$$
$$\hat{cn}_{it+1} = \beta_{20t+1} + \beta_{21t+1} \times \hat{er}_{it+1} + \beta_{22t+1} \times \hat{rn}_{it+1} + \varepsilon_{2it+1}$$
$$\hat{rn}_{it+1} = \beta_{30t+1} + \beta_{31t+1} \times \hat{er}_{it+1} + \beta_{32t+1} \times \hat{cn}_{it+1} + \varepsilon_{3it+1}$$

The regression coefficients relating to the measurement error variance of $\hat{er}_{it+1}$ are:

$$\delta^{M}_{1t+1} = \sigma^2(\upsilon_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} - \{\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})\}$$

$\sigma^2(\upsilon_{1it+1})$ is the measurement error variance of the expected return proxy, $\sigma(er_{it+1}, cn_{it+1})$ $(\sigma(er_{it+1}, rn_{it+1}))$ is the covariance between true expected returns and true cash flow news (return news), and $\sigma(\upsilon_{1it+1}, cn_{it+1})$ $(\sigma(\upsilon_{1it+1}, rn_{it+1}))$ is the covariance between the measurement error in the expected return proxy and true cash flow news (return news).

measurement error variance (i.e., $\delta_1{}^M$) when portfolios are formed on the basis of the book-to-market ratio.

Long-Term Growth Partitions

First, for each annual cross-section of data we partition observations into three groups: (1) bottom third of the distribution of ltg_i (low ltg_i), (2) middle third of the distribution of ltg_i, (medium ltg_i), and (3) top third of the distribution of ltg_i, (high ltg_i). Second, we calculate descriptive statistics for each of the ltg_i partitions for ltg_i and the absolute value of the difference between firm *i*'s actual return on equity in year *t*+2 and the forecast as of year *t* of firm *i*'s year *t*+2 return on equity ($|FE\_ROE_i|$). As shown in Panel A of Table 9, $|FE\_ROE_i|$ is increasing in ltg_i. For example, the mean of $|FE\_ROE_i|$ for observations in the low ltg_i is 5 (7) percent lower than the mean of $|FE\_ROE_i|$ for observations in the medium (high) ltg_i partition, and the untabulated t-statistic for this difference is −6.26 (−8.95).[35] These results are consistent with those shown in Frankel and Lee (1998) and La Porta (1996), and imply that ltg_i is a reasonable *ex ante* proxy for *ex post* forecast errors.

Second, for each proxy we estimate α_1 and $\delta_1{}^M$ via Fama and MacBeth (1973) regressions for each of the three ltg_i partitions. Taken together with the descriptive statistics shown on Panel A of Table 9, the regression results shown on Panel B of Table 9, imply that there is a positive association between the reliability of the expected return proxies and the quality of analysts' earnings forecasts. In particular, as ltg_i increases, α_1 declines and $\delta_1{}^M$ increases for four of the seven proxies (r_{pe}, r_{mpeg}, $r_{\Delta agr}$, and r_{ct}). Moreover, estimates of α_1 for firm-years with low ltg_i are positive for all of the proxies except r_{gls}. It is also pertinent that, for the low ltg_i partition, the estimate of α_1 corresponding to r_{ct} is statistically greater than zero (t-statistic of 3.04) and r_{ct} has a statistically lower measurement error variance (i.e., $\delta_1{}^M$)

---

[35] The mean of $|FE\_ROE_i|$ for observations in the medium ltg_i partition is 2 percent lower than the mean of $|FE\_ROE_i|$ for observations in the high ltg_i partition, and the untabulated t-statistic for this difference is −3.18. Moreover, as shown in Panel A of Table 9, $|FE\_ROE_i|$ is monotonically increasing in ltg_i for all percentiles of $|FE\_ROE_i|$, and all of the t-statistics (untabulated) associated with these differences are greater than 3 in absolute value with the exception of the t-statistic pertaining to the difference between the 95th percentile of $|FE\_ROE_i|$ for firms in the medium ltg_i partition and the 95th percentile of $|FE\_ROE_i|$ for firms in the high ltg_i partition. The t-statistics equal the mean of the annual differences divided by their standard error. Finally, the untabulated Pearson (Spearman) correlation between ltg_i and $|FE\_ROE_i|$ is 0.13 (0.25) with a t-statistic of 6.41 (17.46).

TABLE 8
Multivariate Analyses for Annual Portfolios Using Averages from Portfolios of 10 Formed on the Basis of Beta, the Standard Deviation of Monthly Returns, Market Capitalization, and the Book-to-Market Ratio

Regressions:

$$r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times \hat{ep}_{it+1} + \alpha_{2t+1} \times \hat{cn}_{it+1} + \alpha_{3t+1} \times \hat{rn}_{it+1} + \varepsilon_{it+1}$$

$$\upsilon^{M}_{Cit+1} = \delta_{0t+1} + \delta^{M}_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}$$

	α_1				δ^M_1			
Model	**BETA Portfolios**	***SD* Portfolios**	***SIZE* Portfolios**	***BM* Portfolios**	**BETA Portfolios**	***SD* Portfolios**	***SIZE* Portfolios**	***BM* Portfolios**
pe	−0.11	0.01	−0.82	−2.34	−0.0001	−0.0001	0.0001	0.0007
	(−0.18)	(0.02)	(−1.60)	(−4.42)	(−0.84)	(−1.27)	(2.03)	(5.95)
peg	−0.76	−1.61	−1.05	−1.31	0.0003	0.0005	0.0003	0.0003
	(−1.37)	(−2.74)	(−2.32)	(−3.34)	(3.29)	(3.77)	(4.99)	(7.32)
mpeg	−0.10	−0.53	−0.88	−1.78	0.0001	0.0001	0.0002	0.0004
	(−0.23)	(−0.76)	(−1.76)	(−3.39)	(1.51)	(1.54)	(3.76)	(6.12)
gm	−0.49	−1.18	−1.00	−1.48	0.0002	0.0003	0.0002	0.0004
	(−1.15)	(−2.09)	(−2.28)	(−3.40)	(2.98)	(2.67)	(5.06)	(5.79)
Δagr	−1.05	−1.88	−1.09	−2.08	0.0002	0.0003	0.0002	0.0003
	(−1.61)	(−2.24)	(−2.07)	(−4.27)	(2.95)	(3.19)	(4.44)	(6.13)
ct	0.48	0.52	−0.74	−0.36	0.0000	0.0000	0.0001	0.0001
	(0.88)	(0.77)	(−1.22)	(−0.50)	(−0.16)	(−0.22)	(1.90)	(1.66)
gls	−0.38	−0.95	−1.02	−1.88	0.0000	0.0001	0.0002	0.0003
	(−0.89)	(−1.66)	(−2.98)	(−5.17)	(0.04)	(1.03)	(2.70)	(4.42)

(continued on next page)

TABLE 8 (continued)

Separate regressions are estimated for each of the 18 annual cross-sections of data. Parameter estimates equal the average of the annual regression coefficients.
t-statistics (in parentheses) equal the ratio of the parameter estimates to their temporal standard errors. R^2 is the mean R^2 from the annual regressions.
$\upsilon_{Cit+1} = r_{it+1} - \hat{er}_{it+1} - \hat{cn}_{it+1} - \hat{rn}_{it+1}$, where $\hat{er}_{it+1}$, $\hat{cn}_{it+1}$, $\hat{rn}_{it+1}$, and r_{it+1} represent the portfolio average of the expected return proxy, the cash flow news proxy, the return news proxy, and realized, continuously compounded return for the year $t+1$. Portfolios of 10 are formed by ranking the variables on the basis of the variable of interest: CAPM beta (*BETA*), standard deviation of returns (*SD*), market capitalization (*SIZE*), and book-to-market (*BM*). All the expected return proxies represent continuously compounded returns.
See Table 1 for further details.

$$\upsilon^{M}_{Cit+1} = \upsilon_{Cit+1} - \{\sigma(\hat{er}_{it+1}, \hat{cn}_{it+1}) + \sigma(\hat{er}_{it+1}, \hat{rn}_{it+1})\} \times \hat{\varepsilon}_{1it+1}$$

$$\hat{\varepsilon}_{1it+1} = \frac{-\varepsilon_{1it+1} - \beta_{10t+1}}{\sigma^2(\varepsilon_{1it+1})},\ \hat{\varepsilon}_{2it+1} = \frac{-\varepsilon_{2it+1} - \beta_{20t+1}}{\sigma^2(\varepsilon_{2it+1})},\ \text{and}\ \hat{\varepsilon}_{3it+1} = \frac{-\varepsilon_{3it+1} - b_{30t+1}}{\sigma^2(\varepsilon_{3it+1})}$$

$\sigma(\hat{er}_{it+1}, \hat{cn}_{it+1})$ $(\sigma(\hat{er}_{it+1}, \hat{rn}_{it+1}))$ is the covariance between the expected return proxy and the cash flow news (return news) proxy. The β, ε, and σ^2 (ε) terms shown above are taken from the following first-stage regressions:

$$\hat{er}_{it+1} = \beta_{10t+1} + \beta_{11t+1} \times \hat{cn}_{it+1} + \beta_{12t+1} \times \hat{rn}_{it+1} + \varepsilon_{1it+1}$$

$$\hat{cn}_{it+1} = \beta_{20t+1} + \beta_{21t+1} \times \hat{er}_{it+1} + \beta_{22t+1} \times \hat{rn}_{it+1} + \varepsilon_{2it+1}$$

$$\hat{rn}_{it+1} = \beta_{30t+1} + \beta_{31t+1} \times \hat{er}_{it+1} + \beta_{32t+1} \times \hat{cn}_{it+1} + \varepsilon_{3it+1}$$

The regression coefficients relating to the measurement error variance of $\hat{er}_{it+1}$ are:

$$\delta^{M}_{1t+1} = \sigma^2(\upsilon_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} - \{\sigma(\upsilon_{1it+1}, cn_{it+1}) + \sigma(\upsilon_{1it+1}, rn_{it+1})\}$$

$\sigma^2(\upsilon_{1it+1})$ is the measurement error variance of the expected return proxy, $\sigma(er_{it+1}, cn_{it+1})$ $(\sigma(er_{it+1}, rn_{it+1}))$ is the covariance between true expected returns and true cash flow news (return news), and $\sigma(\upsilon_{1it+1}, cn_{it+1})$ $(\sigma(\upsilon_{1it+1}, rn_{it+1}))$ is the covariance between the measurement error in the expected return proxy and true cash flow news (return news).

than r_{pe} (untabulated t-statistic of 2.29). Hence, for a nontrivial subset of our sample, r_{ct} appears to be a reliable proxy for expected returns.

Finally, in order to provide further insights regarding the association between analysts' forecast errors and the reliability of the expected return proxies, we partition the data into thirds on the basis of $|FE\_ROE_t|$ and estimate α_1 and $\delta_1{}^M$ for each partition. The results of this analysis, which are shown on Panel C of Table 9, support the notion that there is a strong, positive association between the quality of analysts' forecasts and the reliability of the proxies. In particular, as $|FE\_ROE_t|$ increases, α_1 declines and $\delta_1{}^M$ increases. Moreover, for firm-years with low $|FE\_ROE_t|$ estimates of α_1 are positive for all of the proxies (the estimates for r_{mpeg}, r_{gm}, and r_{ct} are statistically significant), and estimates of $\delta_1{}^M$ are not statistically different from zero. While $|FE\_ROE_t|$ is unobservable at time t and, thus, cannot

TABLE 9
Associations between the Long-Term Earnings Growth Rate (ltg_t), Absolute Forecast Errors ($|FE\_ROE_t|$), and Reliability of Expected Return Proxies

Panel A: Descriptive Statistics of ltg_t and $|FE\_ROE_t|$ by ltg_t Partitions

	Mean	Std	5th	25th	50th	75th	95th
ltg_t							
Low ltg_t	0.085	0.024	0.037	0.073	0.093	0.100	0.120
Medium ltg_t	0.138	0.021	0.110	0.120	0.135	0.150	0.180
High ltg_t	0.241	0.086	0.150	0.180	0.215	0.280	0.400
$\|FE\_ROE_t\|$							
Low ltg_t	0.103	0.259	0.003	0.016	0.043	0.106	0.328
Medium ltg_t	0.153	0.388	0.006	0.027	0.063	0.142	0.487
High ltg_t	0.173	0.378	0.010	0.041	0.089	0.179	0.483

Panel B: Regression Analyses by ltg_t Partitions

$$r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times \hat{er}_{it+1} + \alpha_{2t+1} \times \hat{cn}_{it+1} + \alpha_{3t+1} \times \hat{rn}_{it+1} + \varepsilon_{it+1}$$
$$v^M_{Cit+1} = \delta_{0t+1} + \delta^M_{1t+1} \times \varepsilon^A_{1it+1} + \delta_{2t+1} \times \varepsilon^A_{2it+1} + \delta_{3t+1} \times \varepsilon^A_{3it+1} + \mu_{it+1}$$

	α_1			δ_1^M		
Model	**Low ltg_t**	**Medium ltg_t**	**High ltg_t**	**Low ltg_t**	**Medium ltg_t**	**High ltg_t**
pe	0.27	−0.50	−1.82	0.0004	0.0008	0.0011
	(0.65)	(−1.74)	(−4.12)	(1.17)	(2.95)	(2.84)
peg	0.16	−0.25	−1.34	0.0009	0.0008	0.0010
	(0.48)	(−0.67)	(−3.40)	(3.22)	(4.13)	(4.93)
mpeg	0.35	−0.24	−1.20	0.0006	0.0008	0.0009
	(0.99)	(−0.62)	(−2.89)	(2.28)	(3.94)	(4.19)
gm	0.24	0.02	−0.94	0.0009	0.0009	0.0009
	(0.81)	(0.04)	(−2.26)	(3.18)	(4.17)	(4.48)
Δagr	0.20	−0.46	−1.44	0.0006	0.0007	0.0008
	(0.61)	(−1.34)	(−3.49)	(2.38)	(5.13)	(5.16)
ct	1.31	0.44	−0.85	0.0001	0.0002	0.0005
	(3.04)	(0.97)	(−1.40)	(0.27)	(0.73)	(1.64)
gls	−0.79	−0.97	−2.43	0.0008	0.0004	0.0000
	(−2.16)	(−4.30)	(−4.66)	(2.30)	(2.00)	(0.02)

(*continued on next page*)

TABLE 9 (continued)

Panel C: Regression Analyses by $|FE\_ROE_t|$ Partitions

$$r_{it+1} = \alpha_{0t+1} + \alpha_{1t+1} \times \hat{er}_{it+1} + \alpha_{2t+1} \times \hat{cn}_{it+1} + \alpha_{3t+1} \times \hat{rn}_{it+1} + \varepsilon_{it+1}$$

$$v^{M}_{Cit+1} = \delta_{0t+1} + \delta^{M}_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t+1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}$$

	α_1			δ^M_1		
Model	**Low $|FE\_ROE_t|$**	**Medium $|FE\_ROE_t|$**	**High $|FE\_ROE_t|$**	**Low $|FE\_ROE_t|$**	**Medium $|FE\_ROE_t|$**	**High $|FE\_ROE_t|$**
pe	0.41	−0.22	−1.19	−0.0002	0.0000	0.0015
	(1.41)	(−0.63)	(−2.30)	(−0.47)	(0.06)	(2.92)
peg	0.57	−0.36	−1.18	0.0000	0.0008	0.0024
	(1.34)	(−1.14)	(−3.19)	(0.00)	(3.13)	(6.69)
mpeg	0.98	−0.05	−0.98	−0.0003	0.0004	0.0021
	(2.96)	(−0.13)	(−2.54)	(−1.53)	(1.62)	(6.17)
gm	0.76	−0.19	−0.93	−0.0001	0.0006	0.0023
	(2.32)	(−0.68)	(−2.66)	(−0.75)	(2.93)	(6.98)
Δagr	0.60	−0.38	−1.18	−0.0001	0.0005	0.0020
	(1.39)	(−1.09)	(−2.69)	(−0.34)	(2.48)	(5.57)
ct	1.25	1.01	−0.43	−0.0003	−0.0003	0.0007
	(4.06)	(2.52)	(−0.76)	(−1.39)	(−1.50)	(2.28)
gls	0.22	−0.93	−1.72	0.0000	0.0004	0.0003
	(0.79)	(−2.88)	(−3.59)	(0.07)	(1.61)	(1.10)

---

ltg_t ($|FE\_ROE_t|$) partitions are formed each year by sorting the data into thirds on the basis of the consensus analyst forecast of the long-term growth in earnings, ltg_t, (the absolute value of the difference between actual ROE_{t+2} and the forecast as of year t of ROE_{t+2}, $|FE\_ROE_t|$). Low ltg_t ($|FE\_ROE_t|$), medium ltg_t ($|FE\_ROE_t|$), and high ltg_t ($|FE\_ROE_t|$) denote observations with values of ltg_t ($|FE\_ROE_t|$) in the bottom, middle, and top third of the distribution, respectively. Mean, Std, 5th, 25th, 50th, 75th, and 95th represent the mean, standard deviation, 5th percentile, 25th percentile, median, 75th percentile, and 95th percentile, respectively. Separate annual cross-sectional regressions are estimated for each partition. Parameter estimates equal the average of the annual regression coefficients. t-statistics (in parentheses) equal the ratio of the parameter estimates to their temporal standard errors. R^2 is the mean R^2 from the annual regressions.
$v_{Cit+1} = r_{it+1} - \hat{er}_{it+1} - \hat{cn}_{it+1} - \hat{rn}_{it+1}$, where $\hat{er}_{it+1}$, $\hat{cn}_{it+1}$, $\hat{rn}_{it+1}$, and r_{it+1} represent the expected return proxy, the cash flow news proxy, the return news proxy and realized, continuously compounded return for the year $t+1$. All the expected return proxies represent continuously compounded returns.
See Table 1 for further details.

$$v^{M}_{Cit+1} = v_{Cit+1} - \{\sigma(\hat{er}_{it+1}; \hat{cn}_{it+1}) + \sigma(\hat{er}_{it+1}, \hat{rn}_{it+1})\} \times \varepsilon^{A}_{1it+1}$$

$$\varepsilon^{A}_{1it+1} = \frac{-\varepsilon_{1it+1} - \beta_{10t+1}}{\sigma^2(\varepsilon_{1it+1})},\ \varepsilon^{A}_{2it+1} = \frac{-\varepsilon_{2it+1} - \beta_{20t+1}}{\sigma^2(\varepsilon_{2it+1})},\ \text{and } \varepsilon^{A}_{3it+1} = \frac{-\varepsilon_{3it+1} - \beta_{30t+1}}{\sigma^2(\varepsilon_{3it+1})}$$

$\sigma(\hat{er}_{it+1}, \hat{cn}_{it+1})$ ($\sigma(\hat{er}_{it+1}, \hat{rn}_{it+1})$) is the covariance between the expected return proxy and the cash flow news (return news) proxy. The β, ε, and $\sigma^2(\varepsilon)$ terms shown above are taken from the following first-stage regressions:

$$\hat{er}_{it+1} = \beta_{10t+1} + \beta_{11t+1} \times \hat{cn}_{it+1} + \beta_{12t+1} \times \hat{rn}_{it+1} + \varepsilon_{1_{it+1}}$$

$$\hat{cn}_{it+1} = \beta_{20t+1} + \beta_{21t+1} \times \hat{er}_{it+1} + \beta_{22t+1} \times \hat{rn}_{it+1} + \varepsilon_{2_{it+1}}$$

$$\hat{rn}_{it+1} = \beta_{30t+1} + \beta_{31t+1} \times \hat{er}_{it+1} + \beta_{32t+1} \times \hat{cn}_{it+1} + \varepsilon_{3_{it+1}}$$

The regression coefficients relating to the measurement error variance of $\hat{er}_{it+1}$ are:

$$\delta^{M}_{1t+1} = \sigma^2(v_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} - \{\sigma(v_{1it+1}, cn_{it+1}) + \sigma(v_{1it+1}, rn_{it+1})\}$$

$\sigma^2(v_{1it+1})$ is the measurement error variance of the expected return proxy, $\sigma(er_{it+1}, cn_{it+1})$ ($\sigma(er_{it+1}, rn_{it+1})$) is the covariance between true expected returns and true cash flow news (return news), and $\sigma(v_{1it+1}, cn_{it+1})$ ($\sigma(v_{1it+1}, rn_{it+1})$) is the covariance between the measurement error in the expected return proxy and true cash flow news (return news).

be used on an *ex ante* basis, these results remain pertinent as they imply that the development of better approaches for forecasting earnings or statistical models of analysts' forecast errors are fruitful research opportunities.

VI. SUMMARY AND CONCLUSIONS

We develop an empirical approach that allows us to evaluate the reliability of an expected return proxy from its association with realized returns even if realized returns are biased and noisy measures of expected returns. Our results suggest that, for the entire cross-section of firms, the seven accounting-based proxies we consider are not reliable measures of expected returns. None of the proxies has a positive association with realized returns even after controlling for the bias and noise in realized returns attributable to contemporaneous information surprises. Moreover, the simplest model, which is based on the least reasonable assumptions *ex ante*, yields an estimate of expected returns that contains no more measurement error than the remaining proxies. These results are robust as they remain descriptive even after we attempt to purge the proxies of their measurement error via the use of instrumental variables and grouping.

Further analyses, however, demonstrate that certain proxies are reliable for nontrivial subsets of the data. First, we show that the reliability of the proxies is decreasing in the magnitude of consensus analysts' long-term earnings growth rate, and that r_{ct} (the proxy inspired by Claus and Thomas [2001]) is a reliable proxy for firms with low consensus long-term growth forecasts. We also demonstrate that when *ex post* analysts' forecasts errors are low, all of the proxies have a positive association with expected returns. Combining these two sets of results with the fact that *ex post* forecast errors are increasing in analysts' long-term growth forecasts leads us to conclude that the consensus long-term earnings growth rate is a useful *ex ante* indicator of reliability.

This study has three implications for future research. First, the approach we describe can be extended and used in other contexts. For example, development of alternative proxies for cash flow news and return news will improve researchers' ability to draw direct insights about the association between an expected return proxy and true expected returns. This, in turn, will ameliorate the need for the more complex analyses of measurement error variances. In addition, the identity of the model of market equilibrium that best characterizes the risk-return trade-off is of significant interest to many, and our approach provides a means of evaluating models of expected returns suggested by asset-pricing theory.

Second, given the general lack of reliability of the proxies we evaluate, the extant evidence in the accounting and finance literatures based on these proxies should be interpreted with caution. In addition, further research on approaches for imputing reliable estimates of expected returns from contemporaneous prices appears warranted. Finally, by demonstrating a link between the apparent lack of reliability of the proxies we evaluate and the quality of analysts' forecasts, we provide direction for future research—development of better approaches for forecasting earnings or statistical models of analysts' forecast errors appear fruitful.

APPENDIX A

Linear Decomposition of Realized Returns and Estimation of ρ

In this appendix we describe Vuolteenaho's (2002) linear return decomposition, which serves as the foundation for our empirical tests. We also describe the statistical procedure we use to estimate ρ.

Vuolteenaho's (2002) Linear Return Decomposition

We begin by noting the following identities:

$$1 \equiv (1 + ROE_{it+1})^{-1} \times \left(\frac{B_{it+1} + D_{it+1}}{B_{it}}\right) \tag{A.1}$$

$$1 \equiv (1 + R_{it+1})^{-1} \times \left(\frac{M_{it+1} + D_{it+1}}{M_{it}}\right). \tag{A.2}$$

In Equations (A.1) and (A.2) ROE_{it+1} equals E_{it+1}/B_{it}, E_{it+1} is earnings at time $t+1$, B_{it} is equity book value at the end of period t, D_{it+1} is dividends paid during period $t+1$, R_{it+1} is stock return for period $t+1$ and M_{it} is equity market value at the end of period t. We assume clean surplus accounting; hence, Equation (A.1) is an identity (i.e., $B_{it+1} + D_{it+1} \equiv B_{it} + E_{it+1}$).

Dividing each side of Equations (A.1) and (A.2) by D_{it}, rearranging, and taking logs we obtain the following expressions for the ratio of equity book value to dividends and the price to dividend ratio:

$$\begin{aligned}(b_{it} - d_{it}) &= \ln(B_{it}) - \ln(D_{it}) \\ &= \ln\left(\frac{B_{it}}{D_{it}}\right) \equiv \ln(e^{b_{it+1} - d_{it+1}} + 1) + \Delta d_{it+1} - roe_{it+1}\end{aligned} \tag{A.3}$$

$$\begin{aligned}(m_{it} - d_{it}) &= \ln(M_{it}) - \ln(D_{it}) \\ &= \ln\left(\frac{M_{it}}{D_{it}}\right) \equiv \ln(e^{m_{it+1} - d_{it+1}} + 1) + \Delta d_{it+1} - r_{it+1}.\end{aligned} \tag{A.4}$$

In Equations (A.3) and (A.4) lower case letters denote natural logs, Δd_{it+1} is the natural log of dividend growth at time $t+1$ (i.e., $d_{it+1} - d_{it} = \ln(D_{it+1}/D_{it})$), roe_{it+1} denotes the natural log of 1 plus ROE_{it+1} and r_{it+1} denotes realized, continuously compounded returns for time $t+1$. Equation (A.4) illustrates that the price to dividend ratio is increasing in future dividend growth but decreasing in the future discount rate. An analogous interpretation holds for Equation (A.3).

While Equations (A.3) and (A.4) are identities, they are also nonlinear (i.e., they are both functions of logged expressions). Hence, they cannot be directly converted into a linear expression for the natural log of the book-to-market ratio, which is necessary for the development of linear decomposition of realized returns. To derive a linear expression for the natural log of the book-to-market ratio, we approximate Equations (A.3) and (A.4) by taking Taylor expansions of $\ln(e^{m_{it+1} - d_{it+1}} + 1)$ and $\ln(e^{b_{it+1} - d_{it+1}} + 1)$ about the point $e^{\lambda \times (b-d) + (1-\lambda) \times (m-d)}$ (λ is a number between 0 and 1 and $m - d$ and $b - d$ denote unconditional means). Next, we subtract the approximation of Equation (A.4) from the approximation of Equation (A.3). This yields the following linear approximation of the natural log of the book-to-market ratio:

$$\begin{aligned}(b_{it} - m_{it}) &\approx \frac{e^{\lambda \times (b-d) + (1-\lambda) \times (m-d)}}{1 + e^{\lambda \times (b-d) + (1-\lambda) \times (m-d)}} \times (b_{it+1} - m_{it+1}) + r_{it+1} - roe_{it+1} \\ &= \rho \times (b_{it+1} - m_{it+1}) + r_{it+1} - roe_{it+1}.\end{aligned} \tag{A.5}$$

In Equation (A.5) ρ is a number slightly less than 1. To understand ρ better it is useful to assume λ equals 0, in which case $\rho = \frac{e^{(m-d)}}{1 + e^{(m-d)}} = \frac{M/D}{1 + M/D} = \frac{M}{D + M}$, which implies that ρ is bounded between 0 and 1, and is increasing in the price-to-dividend ratio.

Assuming $\lim\limits_{T\to\infty} \rho^{T+1} \times (b_{it+T} - m_{it+T}) = 0$ and iterating Equation (A.5) forward, yields the following approximate expression for the natural log of the book-to-market ratio:

$$(b_{it} - m_{it}) \approx \sum_{\tau=1}^{\infty} \rho^{\tau-1} \times r_{it+\tau} - \sum_{\tau=1}^{\infty} \rho^{\tau-1} \times roe_{it+\tau}. \tag{A.6}$$

Equation (A.6) illustrates that the book-to-market ratio is increasing in future discount rates (i.e., prices are low when expected future discount rates are high) and decreasing in future accounting returns (i.e., prices are high when expected profitability is high).

Finally, we obtain a linear approximation of realized returns by evaluating the change in expectations of Equation (A.6) from t to $t+1$ and rearranging:[36]

$$\begin{aligned} r_{it+1} &\approx E_t[r_{it+1}] + \Delta E_{t+1}\left[\sum_{\tau=1}^{\infty} \rho^{\tau-1} \times roe_{it+\tau}\right] - \Delta E_{t+1}\left[\sum_{\lambda=2}^{\infty} \rho^{\lambda-1} \times r_{it+\lambda}\right] \\ &= er_{it+1} + cn_{it+1} - rn_{it+1}. \end{aligned} \tag{A.7}$$

In Equation (A.7) $E_t[.]$ is the conditional expectation operator and $\Delta E_{t+1}[.]$ equals $E_{t+1}[.] - E_t[.]$. Equation (A.7) illustrates the components of realized returns. In particular, changes in expected future *roe* (i.e., cash flow news, cn_{it+1}) and changes in expected discount rates (i.e., return news, rn_{it+1}) cause realized returns, r_{it+1}, to differ from expected returns, er_{it+1}. Hence, in order to draw meaningful inferences about the reliability of a particular measure of expected returns, we must control for cash flow news and return news.

Estimation of ρ

We estimate ρ for five portfolios—four equal-size portfolios are formed on the magnitude of the price-to-dividend ratio and a fifth portfolio includes all non-dividend-paying stocks. Because ρ is a function of the unconditional (i.e., long-run) price-to-dividend ratio, we set it equal to the median of the annual estimates of ρ_t obtained from the following pooled, cross-sectional regression equation, which is obtained by rearranging Equation (A.5):

$$(b_{i\tau} - m_{i\tau} - r_{i\tau+1} + roe_{i\tau+1}) = \rho_{0t} + \rho_t \times (b_{i\tau+1} - m_{i\tau+1}). \tag{A.8}$$

In Equation (A.8) τ is a number between t and $t-9$ and the sample includes all firm-years in the same price-to-dividend portfolio with requisite data in year $t-9$ through year t (i.e., for each year t we estimate ρ_t using firm years between year $t-9$ and year t).

---

[36] Equation (A.6) implies that: $0 = \Delta E_{t+1}[b_{it} - m_{it}] = \Delta E_{t+1}\left[\sum_{\tau=1}^{\infty} \rho^{\tau-1} \times roe_{it+\tau}\right] - \Delta E_{t+1}\left[\sum_{\lambda=1}^{\infty} \rho^{\lambda-1} \times r_{it+\lambda}\right]$, which equals $\Delta E_{t+1}\left[\sum_{\tau=1}^{\infty} \rho^{\tau-1} \times roe_{it+\tau}\right] - \left\{(r_{it+1} - E_t[r_{it+1}]) + \Delta E_{t+1}\left[\sum_{\lambda=2}^{\infty} \rho^{\lambda-1} \times r_{it+\lambda}\right]\Big]\right\}$. This, in turn, implies Equation (A7). It also implies that cash flow news relates to changes in expectations about all future *roe* (i.e., $t+1$, $t+2$,...), whereas return news relates only to changes in expectations about discount rates pertaining to years occurring *after* year $t+1$.

The values of ρ for the different price-to-dividend portfolios are: (1) non-dividend-paying stocks $\rho = 0.988$, (2) fourth quartile of price-to-dividend ratio for dividend paying stocks (i.e., the quartile with the highest price-to-dividend ratio): $\rho = 0.957$, (3) third quartile of price-to-dividend ratio for dividend-paying stocks: $\rho = 0.921$, (4) second quartile of price to dividend for dividend-paying stocks: $\rho = 0.927$, and (5) first quartile of price-to-dividend ratio for dividend-paying stocks: $\rho = 0.924$.

APPENDIX B

Statistical Approach for Estimating Measurement Error Variances

As discussed in Section II of the main text, the bias in the estimates of α_{1t+1} obtained from Equation (2) does not provide clear evidence about the relative measurement error in a particular expected return proxy. To understand better the bias in α_{1t+1} it is helpful to consider the following re-expression of Equation (1) from the main text:

$$\begin{aligned} r_{it+1} &\approx er_{it+1} + cn_{it+1} - rn_{it+1} \\ &= (e\hat{r}_{it+1} - \upsilon_{1it+1}) + (c\hat{n}_{it+1} - \upsilon_{2it+1}) + (r\hat{n}_{it+1} - \upsilon_{3it+1}) \\ &= e\hat{r}_{it+1} + c\hat{n}_{it+1} + r\hat{n}_{it+1} + (-\upsilon_{1it+1} - \upsilon_{2it+1} - \upsilon_{3it+1}) \\ &= \hat{r}_{it+1} + \upsilon_{Cit+1}. \end{aligned} \qquad \text{(B.1)}$$

In Equation (B.1) υ_{1it+1}, υ_{2it+1}, and υ_{3it+1} denote the measurement error in $e\hat{r}_{it+1}$, $c\hat{n}_{it+1}$, and $r\hat{n}_{it+1}$, respectively. Recall, $r\hat{n}_{it+1}$, which we refer to as the return news proxy, is the negative of return news (i.e., $-rn_{it+1}$). We assume the measurement error in a particular proxy is uncorrelated with the true underlying construct, but may be correlated with the true values of the other constructs and the measurement errors in the remaining proxies. While we are primarily interested in υ_{1it+1}, it is unobservable. However, we can use the combined measurement error, υ_{Cit+1}, to evaluate the extent to which υ_{1it+1} contributes to the bias in α_{1t+1}. Results presented in Rao (1973), Garber and Klepper (1980), and Barth (1991) demonstrate that υ_{Cit+1} and the bias in α_{1t+1} are related in the following manner:

$$\begin{aligned} \upsilon_{Cit+1} &= \alpha_{0t+1} - \mathit{Bias}\,(\alpha_{1t+1}) \times e\hat{r}_{it+1} - \mathit{Bias}\,(\alpha_{2t+1}) \times c\hat{n}_{it+1} - \mathit{Bias}\,(\alpha_{3t+1}) \\ &\quad \times r\hat{n}_{it+1} + \varepsilon_{it+1} \\ &= \alpha_{0t+1} + \left(\frac{\eta_{1t+1}}{\sigma^2\,(\varepsilon_{1it+1})} - \frac{\beta_{21t+1} \times \eta_{2t+1}}{\sigma^2\,(\varepsilon_{2it+1})} - \frac{\beta_{31t+1} \times \eta_{3t+1}}{\sigma^2\,(\varepsilon_{3it+1})}\right) \times e\hat{r}_{it+1} \\ &\quad + \left(\frac{\eta_{2t+1}}{\sigma^2\,(\varepsilon_{2it+1})} - \frac{\beta_{11t+1} \times \eta_{1t+1}}{\sigma^2\,(\varepsilon_{1it+1})} - \frac{\beta_{32t+1} \times \eta_{3t+1}}{\sigma^2\,(\varepsilon_{3it+1})}\right) \times c\hat{n}_{it+1} \\ &\quad + \left(\frac{\eta_{3t+1}}{\sigma^2\,(\varepsilon_{3it+1})} - \frac{\beta_{12t+1} \times \eta_{1t+1}}{\sigma^2\,(\varepsilon_{1it+1})} - \frac{\beta_{22t+1} \times \eta_{2t+1}}{\sigma^2\,(\varepsilon_{2it+1})}\right) \times r\hat{n}_{it+1} + \varepsilon_{it+1}. \end{aligned} \qquad \text{(B.2)}$$

In Equation (B.2) the β and $\sigma^2\,(\varepsilon)$ are the slope coefficients and residual variances obtained from the set of regressions shown in Equation (B.3):

$$\begin{aligned} e\hat{r}_{it+1} &= \beta_{10t+1} + \beta_{11t+1} \times c\hat{n}_{it+1} + \beta_{12t+1} \times r\hat{n}_{it+1} + \varepsilon_{1it+1} \\ c\hat{n}_{it+1} &= \beta_{20t+1} + \beta_{21t+1} \times e\hat{r}_{it+1} + \beta_{22t+1} \times r\hat{n}_{it+1} + \varepsilon_{2it+1} \\ r\hat{n}_{it+1} &= \beta_{30t+1} + \beta_{31t+1} \times e\hat{r}_{it+1} + \beta_{32t+1} \times c\hat{n}_{it+1} + \varepsilon_{3it+1}. \end{aligned} \qquad \text{(B.3)}$$

η_{1t+1}, η_{2t+1}, and η_{3t+1} are functions of the variances of the measurement errors in $\widehat{er}_{it+1}$, $\widehat{cn}_{it+1}$, and $\widehat{rn}_{it+1}$, respectively. Hence, we refer to the η terms as "noise variables." The relation between the noise variables and the covariance structure of the measurement errors is:

$$
\begin{aligned}
-\eta_{1t+1} &= \sigma^2(\upsilon_{1it+1}) + \{\sigma(\upsilon_{1it+1}, \upsilon_{2it+1}) + \sigma(\upsilon_{1it+1}, \upsilon_{3it+2})\} \\
&\quad + \{\sigma(er_{it+1}, \upsilon_{2it+1}) + \sigma(er_{it+1}, \upsilon_{3it+1})\} \\
-\eta_{2t+1} &= \sigma^2(\upsilon_{2it+1}) + \{\sigma(\upsilon_{2it+1}, \upsilon_{1it+1}) + \sigma(\upsilon_{2it+1}, \upsilon_{3it+1})\} \\
&\quad + \{\sigma(cn_{it+1}, \upsilon_{1it+1}) + \sigma(cn_{it+1}, \upsilon_{3it+1})\} \\
-\eta_{3t+1} &= \sigma^2(\upsilon_{3it+1}) + \{\sigma(\upsilon_{3it+1}, \upsilon_{1it+1}) + \sigma(\upsilon_{3it+1}, \upsilon_{2it+1})\} \\
&\quad + \{\sigma(rn_{it+1}, \upsilon_{1it+1}) + \sigma(rn_{it+1}, \upsilon_{2it+1})\}.
\end{aligned}
\tag{B.4}
$$

In Equation (B.4) $\sigma^2(\upsilon_{1it+1})$, $\sigma^2(\upsilon_{2it+1})$, and $\sigma^2(\upsilon_{3it+1})$ denote the variance of the measurement error in $\widehat{er}_{it+1}$, $\widehat{cn}_{it+1}$, and $\widehat{rn}_{it+1}$, respectively. The terms in the first set of braces correspond to covariances between the measurement errors in the proxies (e.g., $\sigma(\upsilon_{1it+1}, \upsilon_{2it+1})$ is the covariance between the measurement error in $\widehat{er}_{it+1}$ and the measurement error in $\widehat{cn}_{it+1}$). The terms in the second set of braces after the equal sign correspond to covariances between true, unobservable constructs and measurement errors in the proxies (e.g., $\sigma(er_{it+1}, \upsilon_{2it+1})$ is the covariance between true expected return and the measurement error in $\widehat{cn}_{it+1}$).

Equations (B.2) through (B.4) demonstrate that the difference between one and α_{1t+1} is a complex function of the covariance structure of the independent variables and the covariance structure of their measurement errors. Hence, this difference is not solely attributable to measurement error in the expected return proxy (i.e., υ_{1it+1}). However, we can use Equations (B.2) through (B.4) to infer the portion of the bias in α_{1t+1} that is attributable to υ_{1it+1}. We follow the two-stage process described in Barth (1991).

In the *first stage* we estimate the regressions shown in Equation (B.3). Next, we combine the regression coefficients, residuals, and residual variances from these regressions to develop a set of constructs that are used as the regressors in the following *second stage* regression, which is obtained by rearranging Equation (B.2):

$$
\begin{aligned}
\upsilon_{Cit+1} &= \eta_{0t+1} + (-\eta_{1t+1}) \times \left(\frac{\varepsilon_{1it+1} - \beta_{10t+1}}{\sigma^2(\varepsilon_{1it+1})}\right) + (-\eta_{2t+1}) \\
&\quad \times \left(\frac{-\upsilon_{2it+1} - \beta_{20t+1}}{\sigma^2(\varepsilon_{2it+1})}\right) + (-\eta_{3t+1}) \times \left(\frac{-\varepsilon_{3it+1} - \beta_{30t+1}}{\sigma^2(\varepsilon_{3it+1})}\right) \\
&\quad + \mu_{it+1} \\
&= \delta_{0t+1} + \delta_{1t+1} \times \varepsilon^{A}_{1it+1} + \delta_{2t+1} \times \varepsilon^{A}_{2it+1} + \delta_{3t-1} \times \varepsilon^{A}_{3it+1} + \mu_{it+1}.
\end{aligned}
\tag{B.5}
$$

Note each of the regressors in Equation (B.5) is a function of a particular error term from one of the regressions shown in Equation (B.3). Hence, we refer to these regressors as adjusted errors (e.g., the regressor that is a function of ε_{1it+1} is referred to as ε^{A}_{1it+1}).

Ranking on the basis of δ_{1t+1} embeds the assumption that $\{\sigma(\upsilon_{1it+1}, \upsilon_{2it+1}) + \sigma(\upsilon_{1it+1}, \upsilon_{3it+1})\}$ and $\{\sigma(er_{it+1}, \upsilon_{2it+1}) + \sigma(er_{it+1}, \upsilon_{3it+1})\}$ are constant across expected return proxies, which may not be descriptive. To circumvent this problem we refine the regression shown in Equation (B.5) by replacing the regressand (i.e., v_{Cit+1}) with v^{M}_{Cit+1}, which is defined as:

$$v^{M}_{Cit+1} = v_{Cit+1} - \{\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})\} \times \varepsilon^{A}_{1it+1}. \tag{B.6}$$

In Equation (B.5) $\sigma(e\hat{r}_{it+1}, c\hat{\ }n_{it+1})$ denotes the covariance between the expected return proxy of interest and the cash flow news proxy, and $\sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})$ is the covariance between the expected return proxy and the return news proxy. The coefficient on ε^{A}_{1it+1} obtained via estimation of Equation (B.6) is:[37]

$$\begin{aligned}\delta^{M}_{1t+1} &= -\eta_{1t+1} - \{\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})\}\\ &= \sigma^{2}(v_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\}\\ &\quad -\{\sigma(v_{1it+1}, cn_{it+1}) + \sigma(v_{1it+1}, rn_{it+1})\}.\end{aligned} \tag{B.7}$$

REFERENCES

Banz, R. 1981. The relation between return and the market value of common stocks. *Journal of Financial Economics* 9: 3–18.

Barth, M. 1991. Relative measurement errors among alternative pension asset and liability measures. *The Accounting Review* 66: 433–463.

Beaver, W. 1970. The time series behavior of earnings. *Journal of Accounting Research* 8: 319–346.

Berk, J., R. Green, and V. Naik. 1999. Optimal investment, growth options, and security returns. *Journal of Finance* 54: 1553–1607.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72: 323–349.

———, and M. Plumlee. 2002. A re-examination of disclosure levels and expected cost of equity capital. *Journal of Accounting Research* 40: 21–40.

———, and ———. 2005. Assessing alternative proxies for the expected risk premium. *The Accounting Review* 80 (January): 21–53.

Brav, A., R. Lehavy, and R. Michaely. 2004. Expected return and asset pricing. Working paper, Duke University.

Campbell, J. 1991. A variance decomposition for stock returns. *The Economic Journal* 101: 157–179.

———, A. Lo, and A. McKinley. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Chen, F., B. Jorgensen, and Y. Yoo. 2004. Implied cost of capital in earnings based valuation: International evidence. *Accounting and Business Research* (December) 34: 323–344.

Claus, J., and J. Thomas. 2001. Equity risk premium as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stocks. *Journal of Finance* 56: 1629–1666.

---

[37] To derive the expression for δ^{M}_{1t+1} we begin by re-expressing the sum of the covariance between the expected return and cash flow news proxies and the covariance between the expected return and return news proxies:

$$\begin{aligned}\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1}) &= \sigma(er_{it+1} + v_{1it+1}, cn_{it+1} + v_{2it+1}) + \sigma(er_{it+1} + v_{1it+1}, rn_{it+1} + v_{3it+1})\\ &= \sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1}) + \sigma(er_{it+1}, v_{2it+1}) + \sigma(er_{it+1}, v_{3it+1})\\ &\quad + \sigma(v_{1it+1}, cn_{it+1}) + \sigma(v_{1it+1}, rn_{it+1}) + \sigma(v_{1it+1}, v_{2it+1}) + \sigma(v_{1it+1}, v_{3it+1}).\end{aligned}$$

Next, using the definition of $-\eta_{1t+1}$ shown in Equation (B.4), we arrive at the following expression for δ^{M}_{1t+1}:

$$\begin{aligned}\delta^{M}_{1t+1} &= -\eta_{1t+1} - \{\sigma(e\hat{r}_{it+1}, c\hat{n}_{it+1}) + \sigma(e\hat{r}_{it+1}, r\hat{n}_{it+1})\}\\ &= \sigma^{2}(v_{1it+1}) - \{\sigma(er_{it+1}, cn_{it+1}) + \sigma(er_{it+1}, rn_{it+1})\} - \{\sigma(v_{1it+1}, cn_{it+1}) + \sigma(v_{1it+1}, rn_{it+1})\}.\end{aligned}$$

Cochrane, J. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Dhaliwal, D., L. Krull, O. Li, and W. Moser. 2004. Dividend taxes and implied cost of capital. Working paper, The University of Arizona.

Easton, P., G. Taylor, P. Shroff, and T Sougiannis. 2002. Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research* 40: 657–676.

———. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79: 73–96.

Elton, J. 1999. Expected return, realized return, and asset pricing tests. *Journal of Finance* 54: 1199–1220.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 81: 607–636.

———, and K. French. 1992. The cross-section of expected returns. *Journal of Finance* 47: 427–465.

———, and ———. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 154–194.

———, and ———. 2002. The equity premium. *Journal of Finance* 57: 637–659.

Francis, J., I. Khurana, and R. Pereira. 2003. Global evidence on incentives for voluntary disclosures and the effect on cost of capital. Working paper, University of Missouri–Columbia.

———, R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79 (October): 967–1010.

Frankel, R., and C. Lee. 1998. Accounting valuation, market expectations, and cross-sectional stock returns. *Journal of Accounting and Economics* 25: 283–319.

Freeman, R., J. Ohlson, and S. Penman. 1982. Book rate-of-return and prediction of earnings changes: An empirical investigation. *Journal of Accounting Research* 20: 3–42.

Garber, S., and S. Klepper. 1980. "Administrative pricing" or competition coupled with errors of measurement? *International Economic Review* 21: 413–435.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Towards an ex-ante cost of capital. *Journal of Accounting Research* 39: 135–176.

Gode, D., and P. Mohanram. 2003. Inferring the cost of capital using the Ohlson-Juettner model. *Review of Accounting Studies* 8: 399–431.

Guay, W., S. Kothari, and S. Shu. 2003. Properties of the implied cost of capital using analysts' forecasts. Working paper, Massachusetts Institute of Technology.

Hail, L. 2002. The impact of voluntary corporate disclosure on the ex ante cost of capital for Swiss firms. *European Accounting Review* 11: 741–773.

———, and C. Leuz. 2004. International differences in cost of capital: Do legal institutions and securities regulation matter? Working paper, University of Pennsylvania.

Hribar, P., and N. Jenkins. 2004 The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies* 9: 337–356.

La Porta, R. 1996. Expectations and the cross-section of stock returns. *Journal of Finance* 51: 1715–1742.

Lee, C., D. Ng, and B. Swaminathan. 2003. International asset pricing: Evidence from the cross-section of implied cost of capital. Working paper, Cornell University.

Lintner, J. 1965. Security prices, risk, and maximal gains from diversification. *Journal of Finance* 20: 587–615.

Malkiel, B., and Y. Xu. 1997. Risk and return revisited. *Journal of Portfolio Management* 23: 9–14.

Mossin, J. 1966. Equilibrium in a capital asset market. *Econometrica*: 768–783.

O'Hanlon, J. and A. Steele. 2000. Estimating the equity risk premium using accounting fundamentals. *Journal of Business Finance and Accounting* 27: 1051–1084.

Ohlson, J., and B. Juettner-Nauroth. 2003. Expected EPS and EPS growth as determinants of value. Working paper, New York University.

Richardson, S., S. Teoh, and P. Wysocki. 2001. Tracking analyst's forecasts over the annual earnings horizon: Are analyst's forecasts optimistic or pessimistic? Working paper, University of Michigan.

Rao, P. 1973. Some notes on the errors-in-variables model. *The American Statistician* 27: 217–218.

Sharpe. W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425–442.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71: 289–315.

Vuolteenaho, T. 2002. What drives firm-level stock returns? *Journal of Finance* 57: 233–264.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 539–561

Restoring Trust after Fraud: Does Corporate Governance Matter?

David B. Farber
Michigan State University

ABSTRACT: In this study, I examine the association between the credibility of the financial reporting system and the quality of governance mechanisms. I use a sample of 87 firms identified by the SEC as fraudulently manipulating their financial statements. Consistent with prior research, results indicate that fraud firms have poor governance relative to a control sample in the year prior to fraud detection. Specifically, fraud firms have fewer numbers and percentages of outside board members, fewer audit committee meetings, fewer financial experts on the audit committee, a smaller percentage of Big 4 auditing firms, and a higher percentage of CEOs who are also chairmen of the board of directors. However, the results indicate that fraud firms take actions to improve their governance, and three years after fraud detection these firms have governance characteristics similar to the control firms in terms of the numbers and percentages of outside members on the board, but exceed the control firms in the number of audit committee meetings. I also investigate whether the improved governance influences informed capital market participants. The results indicate that analyst following and institutional holdings do not increase in fraud firms, suggesting that credibility was still a problem for these firms. However, the results also indicate that firms that take actions to improve governance have superior stock price performance, even after controlling for earnings performance. This suggests that investors appear to value governance improvements.

Keywords: *fraud; corporate governance; credible financial reporting; investor trust; agency costs; independent directors; audit committee.*

Data Availability: *The data used in this study are available from public sources identified in the text.*

This paper is based on my dissertation completed at Cornell University. I thank my dissertation committee chairperson, Julia D'Souza, for her unwavering support and guidance in the development of this paper. I also thank my other committee members Charles Lee, Tim Mount, and Bhaskaran Swaminathan, as well as Daniel Beneish, Walt Blacconiere, Tom Dyckman, John Elliott, Sue Haka, Marilyn Johnson, Kathy Petroni, and workshop participants at the University of California at Berkeley, Cornell University, Georgia State University, Indiana University, Michigan State University, University of Nebraska, New York University, and Syracuse University for their suggestions and comments. I appreciate the valuable comments I received at the 2002 AAA Annual Meeting. I thank Mark Beasley for sharing his sample of fraud firms. I am also grateful to the librarians at Cornell University and the University of Rochester for their invaluable assistance, with special thanks to Don Schnedeker. I thank I/B/E/S for analyst data. I thank the editor (Patricia Dechow) and two anonymous referees for helpful suggestions that have greatly improved this paper. Finally, I acknowledge the financial support of the Johnson Graduate School of Management. All errors are my own.

Editor's Note: This paper was accepted by Patricia Dechow, Editor.

Submitted January 2003
Accepted September 2004

I. INTRODUCTION

Empirical evidence indicates that weak corporate governance is associated with financial reporting fraud (e.g., Dechow et al. 1996 [hereafter DSS]; Beasley 1996), but little is known about the actions that fraud firms take to improve their weak governance after fraud detection and, perhaps more importantly, how effectively these actions restore investor trust. Given the importance of the relation between the quality of governance mechanisms and the credibility of the financial reporting system, it is surprising that we know so little about the nature and extent of this relation. In this study, I seek to fill this gap by investigating the association between the revelation of financial reporting fraud and subsequent improvements in the quality of corporate governance mechanisms and the corresponding economic consequences of such improvements. This evidence is important because the quality of corporate governance is at the forefront of concerns of regulators and corporate management.

Regulators, perhaps in response to the recent flurry of highly publicized financial reporting frauds (e.g., Enron), are considering or have enacted rules intended to strengthen the quality of corporate governance.[1] Much of the debate surrounding these proposals revolves around the idea that stronger governance is associated with more credible financial reporting. Some firms accused of fraud are using a recovery strategy that includes governance improvements. For example, in a recent financial press article (Sorkin 2002), Edward D. Breen, who was appointed CEO of Tyco following allegations of fraudulent financial reporting there, commented on the state of the quality of corporate governance at Tyco:

> The most important thing, in my opinion, when I came into the company, and I've stated this many times, was *fixing* corporate governance ... (emphasis added)

Breen appears to believe that there is a net benefit to improving his firm's governance mechanisms, perhaps in the form of a positive stock price reaction.

It seems obvious that firms caught committing fraud would take some actions with respect to their governance mechanisms, the most likely of which is to discharge the culprits associated with the fraud. However, it is not as obvious that these firms would necessarily *improve* their governance after fraud detection because they could merely replace inside directors with other insiders, rather than with outside directors, for example.

My first main research objective, therefore, is to examine whether there is an association between the detection of financial reporting fraud and subsequent improvements in the quality of governance mechanisms. I find that while fraud firms' governance structures are initially weak, by the end of a three-year period following the year of fraud detection these firms have a similar board of director profile and proportion of firms with the combined CEO/COB position to that of their matched control firms. I also find that fraud firms hold more audit committee meetings than their matched control firms by the end of this same period. Overall, these results suggest that improving the quality of governance is important to fraud firms.

Given that regulators are mandating new governance rules and that fraud firms are likely to be expending scarce resources on governance improvements after fraud detection, it is important to document whether these improvements provide *any* economic benefits. Evidence that governance improvements do indeed provide economic benefits would support the basis for these rules and firms' expenditures on enhanced governance. However, it could very well be the case that these improvements do not provide any economic benefits and are merely window dressing (i.e., non-value-enhancing) and thus an inefficient use of resources. Providing evidence on either case adds to our knowledge of the economics of improving the quality of the corporate governance mechanisms that monitor the financial reporting process.

---

[1] For example, NYSE (2004) and NASDAQ (2004) have both recently enacted rules requiring that their member firms have boards of directors with a majority of independent directors.

My second main research objective, therefore, is to examine whether these improvements aid fraud firms in restoring their reputations with informed capital market participants and recovering share values. Specifically, I examine the relation between improvements that fraud firms make in their boards and audit committees following fraud detection and the corresponding responses of analysts, institutional investors, short sellers, and the stock market. Results indicate that analyst following and institutional ownership do not increase in fraud firms, suggesting that credibility is likely still a problem for these firms. These results are not surprising given that fraud firms experience negative returns following fraud detection. More interestingly, I find a positive and economically significant relation between increases in board independence and long-run buy-and-hold abnormal returns over the three-year period following fraud detection, even after controlling for earnings performance. Overall, my findings suggest that fraud firms and, perhaps more importantly, the market, view improving the quality of governance mechanisms as a way of restoring trust after fraud.

The remainder of this paper is organized as follows. In Section II, I provide a brief review of the relevant literature on governance mechanisms and fraud. Section III contains a description of the sample. Section IV contains the research design and empirical results. I summarize and conclude in Section V.

II. LITERATURE REVIEW

One stream of the accounting literature related to financial reporting fraud examines the association of governance mechanisms and the likelihood of fraud (e.g., DSS; Beasley 1996). DSS find that firms committing financial reporting fraud are more likely to have a board of directors dominated by insiders and are less likely to have an audit committee. However, they do not find that the use of a (then) Big 6 audit firm significantly influences the likelihood of fraud. Beasley (1996) finds that the inclusion of larger proportions of outside members on the board of directors significantly reduces the likelihood of financial reporting fraud, but does not find significant evidence that fraud firms' audit committees meet less frequently. In contrast to DSS and Beasley (1996), the current study investigates the magnitude and economic consequences of fraud firms' improvements in important monitoring mechanisms during the three-year period following fraud detection.

There are few extant studies that examine firm *responses* to fraud detection. In research closely related to the current study, Agrawal et al. (1999) investigate the association of fraud detection and subsequent turnover in senior managers and directors, but fail to find one. Agrawal et al. (1999) examine 103 firms that commit a broad range of frauds, but only four of those firms represent cases of financial reporting fraud. Thus, it is difficult to draw inferences from their study regarding the effect of GAAP violations on governance improvements. Moreover, past research (Karpoff and Lott 1993) finds that mean abnormal returns in a two- or three-day window around fraud detection are significantly more negative in cases of financial reporting fraud than in other cases of fraud. Therefore, incentives to overhaul governance mechanisms are likely to be greater for cases of financial reporting fraud firms than for other cases of fraud. However, this remains an empirical question—one that I examine in this study.

Livingston (1997) and Beneish (1999) investigate senior management turnover following the detection of financial reporting fraud. Srinivasan (2004) investigates outside director turnover after an accounting restatement. Livingston (1997) finds significant turnover in top managers and financial officers, while Beneish (1999) finds no unusual turnover in senior management. Srinivasan (2004) finds significantly higher turnover in boards for firms that restate earnings downward and that the likelihood of director departure increases in restatement severity. Unlike the Livingston (1997) and Beneish (1999) studies, the current study focuses on the mechanisms intended to oversee managerial actions, rather than on management itself. This is important because agency costs will continue to exist as long there is separation of ownership and control in a firm. Replacing top management following fraud detection does not get to the root of the problem. As Jensen (1993) notes, "... bad systems or rules, not bad people, underlie the general failings of the board of directors." Therefore, enhancing the quality of governance mechanisms is more likely to ensure that managers do not deviate from their fiduciary duties. Unlike the Srinivasan (2004) study, the current study focuses on changes in a more comprehensive set of governance variables. The current study also examines these changes around a financial reporting fraud, a distinctly different event from an earnings restatement, which does not necessarily involve fraud. Moreover, the current study also investigates the economic consequences of these governance changes for the firms themselves.

The only extant study of which I am aware that provides direct evidence on the market's response to *changes* in boards of directors is that by Rosenstein and Wyatt (1990), who, in a nonfraud setting, find a small positive mean abnormal return over a two-day window centered on the announcement dates of outside director appointments. It is possible, however, that the market will discount or ignore governance changes made by fraud firms because the revelation of fraud has so badly tarnished these firms' reputations. Therefore, our understanding of the economic consequences of governance changes after fraud detection is an open empirical issue that warrants investigation.

III. SAMPLE SELECTION

Fraud Sample

The fraud sample consists of publicly held companies cited in United States Securities and Exchange Commission (SEC) Accounting and Auditing Enforcement Releases (AAERs)[2] for violating SEC Rule 10b-5 during the period 1982–2000. An illustrative example of the SEC's justification for issuing AAERs under the fraud provisions of the Exchange Act of 1934 follows (per AAER No. 689, in the case of *Kurzweil Applied Intelligence, Inc.*):

> Section 10(b) of the antifraud provisions of the Exchange Act (of 1934) and Rule 10b-5 thereunder proscribes the making of materially false and misleading statements "in connection with the purchase or sale of any security." Violations of Section 10(b) and Rule 10b-5 occur when an issuer makes material misstatements in registration statements, prospectuses or periodic reports filed with the Commission and trading thereafter occurs in the issuer's securities. The filing of false and misleading reports is also a violation of Section 10(b) of the Exchange Act (of 1934) and Rule 10b-5 thereunder because reports of publicly traded companies affect the markets for offer, sale and purchase of their securities.

---

[2] The SEC issues AAERs against firms and others (e.g., audit firms, managers, and directors) for violations of SEC rules.

I use AAERs as a proxy for the occurrence of fraud. Consistent with Beasley (1996) and Bonner et al. (1998), this proxy is intended to capture extreme cases of fraud in the error-to-fraud continuum. A focus on AAERs is justified because while the SEC cannot possibly investigate every case of fraud, it does target the highest profile cases (Feroz et al. 1991). Consequently, the fraud sample used in this study is likely to include the most egregious cases of fraud and thus provides a powerful setting to examine the impact of fraud detection on governance improvements because if no effect is found in this setting, then it is not likely to be found in any other. However, the use of AAERs has limitations. For example, because the SEC selects cases for which it has the best chance of winning a judgment, they are likely to include instances of the most extreme misleading reporting. Therefore, the results of this study are not likely to be generalizable to the entire population of firms that report fraudulently.

I restrict my analyses to frauds that were *detected* through 1997 to allow for a three-year period after fraud detection in which to examine governance changes. The analysis period is restricted to five years to make data collection more tractable. Table 1, Panel A reconciles the number of AAERs issued over my analysis period to my final fraud sample. There were 1,357 AAERs issued from 1982 to 2000,[3] of which 1,067 did not involve financial statement fraud, were duplicates, or were outside the scope of the analysis period. I was unable to locate proxies or financial statement data for another 195 firms. I include in the final fraud sample each fraud firm that has either proxy statements or SEC Forms 10-K available for at least one year before and at least one year after fraud detection. This sample selection criteria initially leads to 95 firms in the fraud sample, but because I was unable to identify matches for 8 of the fraud firms, the final fraud sample size is 87. Once I identify a firm for inclusion in the fraud sample, I search the Dow Jones database for the period surrounding the fraud period identified in the AAER to determine the fraud detection date. Following Feroz et al. (1991), if no indication of the fraud detection date appears in Dow Jones, I use the AAER date as a proxy for the detection date. During my review of the AAERs, I carefully examined each one to ensure that the fraud firms identified were cited for false and misleading financial reports. Moreover, I noted no cases where a firm was cited for violating Rule 10b-5 only for reasons other than the issuance of false or misleading financial statements.

Panel B of Table 1 indicates that fraud firms are widely distributed among industries, with some clustering of firms in machinery and computer equipment, measurement instruments, and business services. Panel C of Table 1 indicates that 60 percent of the frauds involve fictitious transactions (i.e., those that are unambiguously fraudulent and include creating phony invoices for sales that do not exist and never will), that fraud firms overstate their net income by a mean (median) of 174 percent (53.70 percent) and that the mean (median) length of the fraud is 6.66 (6.00) quarters.

Control Sample

Prior research (e.g., DSS) and the current study show significant clustering by industry among fraud firms. Therefore, examination of governance changes surrounding fraud detection is likely to reveal characteristic governance behavior that is associated with changes in the industry rather than with fraud detection. To control for this possibility, I use a matched-pairs design to test changes in corporate governance mechanisms following fraud detection.

---

[3] Because AAERs are typically issued two to three years subsequent to fraud detection (see, Feroz et al. 1991), I reviewed them through the year 2000 to ensure that I identified all frauds through the year 1997 (AAERs contain information regarding the fraud period).

For each fraud firm, I select a control firm with the same four-digit SIC code, with net sales within ±25 percent of the fraud firm's net sales for the year preceding fraud detection,

TABLE 1
Fraud Sample Selection, Fraud Firm Characteristics, and Fraud-Specific Statistics

Panel A: Sample Selection of 87 Firms Subject to Enforcement Actions under SEC Rule 10b-5 between 1982 and 2000

Number of AAERs issued between 1982 and 2000	1,357
Less:	
AAERs not involving financial statement fraud, duplicate AAERs, or AAERs outside scope of analysis period (e.g., unintentional misapplication of GAAP, AAERs against auditors, AAERs issued for frauds occurring after 1997)	1,067
Firms without proxy or financial statement data	195
Firms without an available matching non-fraud firm	8
Final Sample	87

Panel B: Distribution of Fraud Firms by SIC Code

Two-Digit SIC Code	Industry Description	Number of Firms
13	Oil and Gas Extraction	5
22	Textile Mill Products	1
23	Apparel	2
26	Paper and Allied Products	1
27	Printing and Publishing	2
28	Chemicals	5
32	Stone, Clay and Glass	1
34	Fabricating Metals	5
35	Machinery and Computer Equipment	10
36	Electrical	9
37	Transportation Equipment	1
38	Measurement Instruments	9
39	Miscellaneous Manufacturing	1
42	Motor Freight Trans, Warehousing	1
45	Air Transportation	1
49	Electric, Gas, Sanitary Services	4
50	Wholesale Durable Goods	5
52	Building Materials	1
56	Apparel and Accessory Stores	2
59	Miscellaneous Retail	4
61	Nondepository Credit Institution	1
73	Business Services	12
79	Amusements and Recreation	1
80	Health Services	2
99	Miscellaneous	1
	Total	87

(continued on next page)

TABLE 1 (continued)

Panel C: Fraud-Specific Statistics

Type of GAAP Violation[a]	Number of Firms		Percentage of Total		
Fictitious Transactions	52		60		
Timing Difference Transactions	31		36		
Unidentified	4		4		
Total	87		100		
	Median	**Mean**	**Std. Deviation**	**Max**	**Min**
Overstatement as percentage of net income	53.70	174.20	361.86	2,169	0
Length of GAAP Violation (in quarters)	6.00	6.66	4.21	19	1

[a] Fictitious transactions are unambiguously fraudulent and include phony invoices for sales that do not exist and never will. Timing difference transactions represent premature recognition of legitimate transactions that eventually occur.

and within the same stock exchange. If I cannot find a match within the same four-digit SIC code, then I relax my matching criteria to include firms with the same three-digit SIC code, and if necessary, match firms on the basis of a two-digit SIC code. The potential control firm is included in the final matched control sample if there are no reports of fraud in the Dow Jones Database for that firm in the two years before and three years after the fraud detection year of its corresponding fraud firm. Additionally, I search AAERs issued during the analysis period to ensure that the SEC has not cited the potential control firm. The final non-fraud sample consists of 87 firms. Table 2 indicates that fraud and non-fraud firms do not differ significantly based on net sales, market value, or exchange listing.[4]

IV. RESEARCH DESIGN AND EMPIRICAL RESULTS

Changes in Governance Variables

Prior research (Beasley 1996; DSS) has established that, compared to matched control firms, fraud firms have weaknesses in several key governance mechanisms. These weaknesses include boards with lower percentages of outside directors, a higher proportion of firms with the combined CEO/COB position, a lower proportion of firms with audit committees, fewer audit committee meetings, and a smaller percentage blockholder ownership. Other governance mechanisms likely to be important to the financial reporting system include the composition of the audit committee, including its percentage of independent directors and financial experts, the quality of the external audit firm, and the percentage of ownership by managers and directors.

Several empirical studies show that boards composed mainly of outside directors are more effective than boards composed mainly of insiders (e.g., Brickley and James 1987; Weisbach 1988; Rosenstein and Wyatt 1990), and that boards are ineffectual monitors when the board's equity ownership is small and when the CEO is also Chairman of the Board (Jensen 1993). Agrawal et al. (1999) provide the only empirical study of which I am aware that investigates changes in both senior management and directors after fraud detection.

---

[4] I also test for the difference in age between fraud and control firms and find no statistical difference.

TABLE 2
Matching Statistics for Fraud and Control Firms

Variable	Fraud Firm Mean (in millions) (n = 87)	Control Firm Mean (in millions) (n = 87)	t-statistic	Fraud Firm Median (in millions) (n = 87)	Control Firm Median (in millions) (n = 87)	p-value
Net Sales	373.61	343.59	0.77	37.07	44.85	0.28
Market Value	336.59	315.94	0.40	55.48	31.39	0.14

Exchange Listing	Fraud Firm	Control Firm
NYSE	17	17
AMEX	5	5
NASDAQ/OTHER	65	65
Total	87	87

Fraud firms are matched with control firms on the basis of year, net sales, SIC code, and stock exchange listing. The t-statistic is for the difference between the means of the matched pairs. The p-value is for the Wilcoxon signed rank test.

They find no significant turnover in either senior management or directors in the three years after fraud detection. However, it is difficult to draw inferences from their study about the detection of GAAP violations as a catalyst for making improvements in governance because their sample included only four GAAP violators.

Sommer (1991) suggests that having an audit committee as part of the governance structure and having an effective audit committee are two different matters; further, the audit committee needs to meet on a regular basis without management in order for it to be effective. He also notes that the audit committee, because it has constant access to the internal audit staff, external audit firm, and other corporate personnel, is likely to be the first to identify a potential irregularity. His study supports the notion that there is a positive relation between the number of meetings an audit committee holds and its effectiveness. Beasley (1996), however, finds no significant difference between fraud and non-fraud firms in the number of audit committee meetings held in the year prior to fraud detection. The Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (BRC) (1999) suggests that the composition of the audit committee is also likely to be critical to its performance. I therefore examine whether the composition of the audit committee differs for fraud and control firms along the dimensions of size, number of outside directors, and number of financial experts.

The extant literature finds no significant difference between fraud and non-fraud firms in audit-firm quality as measured by the use of Big 4 auditing firms versus non-Big 4 firms in the year prior to fraud detection (DSS; Beneish 1997). Several studies find an association between audit-firm quality and audit-firm size (e.g., DeAngelo 1981). One implication of this finding is that a larger audit firm is likely to be a stronger monitor than a smaller one. Palmrose (1988) suggests that (then) Big 8 audit firms are sued less frequently than other audit firms because they provide higher quality audits. DeFond (1992) finds a positive relation between audit-firm quality, as proxied by audit-firm size and name brand (i.e., Big 4), and its ability to mitigate agency problems. In sum, these findings suggest that Big 4 audit firms provide better monitoring than other categories of audit firms.

Firms choosing governance mechanisms likely also evaluate their costs. Empirical studies find that internal control improvements are costly to firms in terms of the time and effort required to improve governance (Yermack 1996; Klein 2002). Moreover, there is a prohibitive cost associated with establishing internal controls that would eliminate entirely the probability of (future) frauds (Jensen 1993). Notwithstanding these costs, because of the severity of the agency problems of the fraud firms used in this study, it is likely that their costs of not improving governance are greater than their costs of doing so. Economic theory suggests that efforts to repair the financial reporting system must necessarily include significant improvements in its monitoring components. More specifically, agency theory suggests that the demand for monitoring is positively related to the prevalence of agency costs (Fama 1980; Fama and Jensen 1983). Because fraud is a manifestation of acute agency costs, I expect that fraud firms will experience a greater increase in: (1) outside director percentage, (2) audit committee activity, (3) audit committee independence, (4) audit committee financial expertise, and (5) audit-firm quality; and a greater decrease in the proportion of firms with the combined CEO/COB position after fraud detection than will their corresponding control firms over the same period. Blockholders are also likely to play a key monitoring role (Jensen 1993), but because they are an external monitoring mechanism, it is difficult to predict changes in their post-fraud detection ownership. It is also difficult to predict the change in inside ownership by management and directors given that there is likely to be substantial turnover in both following the detection of fraud. I therefore do not make predictions for changes in the holdings of blockholders, managers, or directors.

I obtain data on these governance variables from the firm's proxy statement or SEC Form 10K. Figure 1 illustrates the direction of the changes in the governance variables for the sample of fraud firms from the year prior to fraud detection compared to the third year afterward. Figure 1, Panel A indicates that increases occur in the percentage of outside directors[5] (*OutsideDir%*), and the percentage of stock held by blockholders (*Block%*), while the percentage of firms with the combined CEO/COB position (*CEO=COB%*) and the percentage of stock held by management and directors (*InsideOwn%*) declines. Panel A also indicates that the percentage of firms with a Big 4 audit firm (*Big4%*) and the percentage of institutional ownership (*InstOwn%*) remain virtually the same.

Figure 1, Panel B provides evidence that fraud firms increase the number of outside directors on their boards (*#OutsideDir*), the number of audit committee meetings (*#AudComMeet*), the number of audit committee members (*#AudComMbrs*), and the number of outside directors on their audit committees (*#AudComOutsideDir*). This figure also indicates that the number of financial experts[6] (*#FinlExperts*) remains virtually unchanged.

Table 3 provides tests of significance of these governance changes using a set of control firms. I first test whether the sample of fraud firms used in this study has a similar governance profile to that used in prior studies. I then test my predictions about changes in these governance variables. As reported in Table 3, Panel A, I find that the mean *OutsideDir%* in the year preceding fraud detection (Initial) is 47 percent for fraud firms compared to 56 percent for control firms, a statistically significant difference (t = –2.54), which is consistent with Beasley (1996) and DSS. Panel A also documents that *#OutsideDir* in Initial is 3.3 for fraud firms compared to 4.1 for control firms, a statistically significant

---

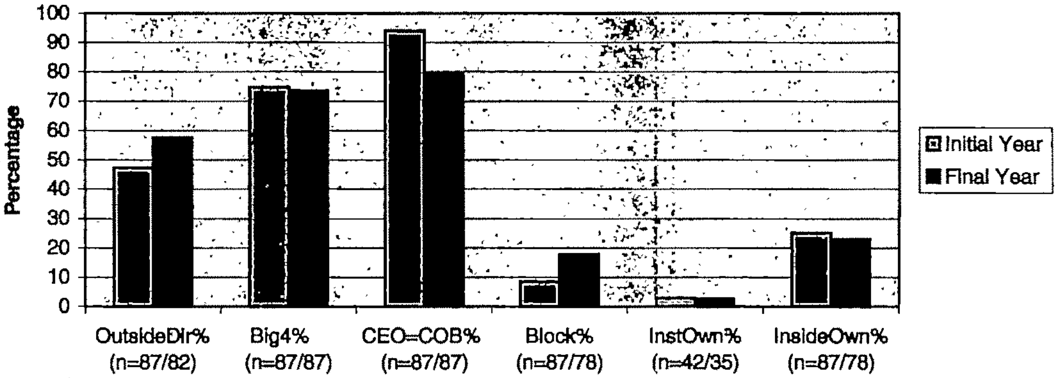
[5] Consistent with Rosenstein and Wyatt (1990), I define an outside director as a director who is not a present or former employee of the firm and whose only formal connection to the firm is his/her duty as a director.

[6] I define a financial expert as one who has accounting or related financial management expertise (i.e., a person who is or has been chief financial officer, or other senior corporate officer with financial oversight responsibility). This definition is consistent with that in the report of the Blue Ribbon Committee (1999).

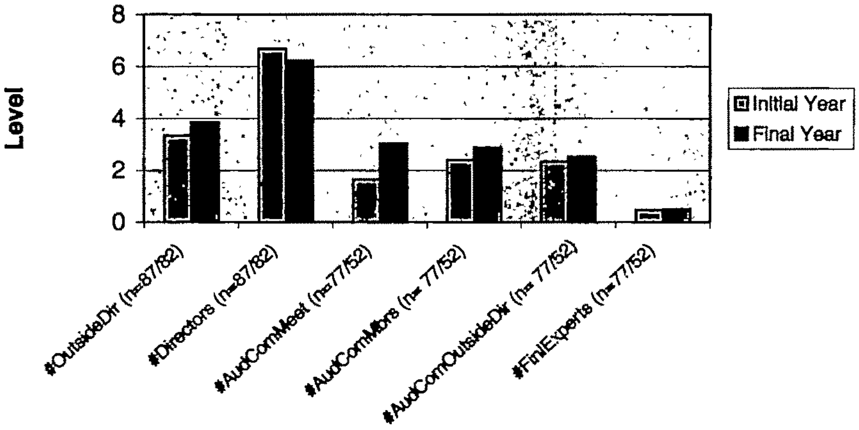
FIGURE 1

Governance Variables for Fraud Firms in the Year Prior to Fraud Detection (Initial) and the Third Year Afterward (Final) for Frauds Detected from 1982 to 2000

Panel A: Means of Governance Variables in Percentage (sample size = Initial/Final)



Panel B: Means of Governance Variables in Levels (sample size = Initial/Final)



OutsideDir% = outside director percentage;
Big4% = percentage with Big 4 audit firms;
CEO=COB% = percentage with the combined CEO/COB position;
Block% = percentage of shares held by > 5 percent blockholders;
InstOwn% = the percentage of shares held by institutions that file SEC Form 13F;
InsideOwn% = the percentage of shares held by management and directors;
#OutsideDir = number of outside directors;
#Directors = number of directors;
#AudComMeet = number of audit committee meetings;
#AudComMbrs = number of audit committee members;
#AudComOutsideDir = number of outside directors on audit committee;
#FinlExperts = number of financial experts on audit committee;
Initial Year = the year prior to fraud detection; and
Final Year = the third year after fraud detection.

difference (t = –2.08). The change in *OutsideDir%* is greater for fraud firms in years 1 through 4, thus providing support for the hypothesis that fraud firms will increase their

TABLE 3
Univariate Comparisons of Board of Director, Audit Committee, and Other Corporate Governance Variables

Panel A: Board of Director Characteristics

		Year					
Variable	**Firm**	**Initial[a]**	**1[b]**	**2[b]**	**3[b]**	**4[b]**	**Final[a]**
OutsideDir%	Fraud	47.03	3.25	3.50	7.21	9.95	55.77
	Control	55.68	−0.99	−1.26	−1.86	1.99	56.61
	(t-stat)	(−2.54)***	(2.11)**	(1.62)*	(2.87)***	(1.96)**	(−0.20)
	(#matched pairs)	(87)	(84)	(85)	(86)	(79)	(79)
#OutsideDir	Fraud	3.34	0.024	0.13	0.31	0.51	3.68
	Control	4.07	−0.04	−0.17	−0.20	−0.38	3.96
	(t-stat)	(−2.08)**	(0.36)	(1.14)	(1.89)**	(1.58)*	(−0.73)
	(#matched pairs)	(87)	(84)	(85)	(86)	(79)	(79)
#Directors	Fraud	6.69	−0.44	−0.45	−0.49	−0.52	6.05
	Control	7.00	0.05	−0.21	−0.27	−0.31	6.63
	(t-stat)	(−1.11)	(−2.41)**	(−0.84)	(−0.73)	(−0.66)	(−1.05)
	(#matched pairs)	(87)	(84)	(85)	(86)	(79)	(79)

Panel B: Audit Committee Characteristics

Variable	**Firm**	**Initial[a]**	**1[b]**	**2[b]**	**3[b]**	**4[b]**	**Final[a]**
#AudComMeet	Fraud	1.61	0.85	1.39	1.38	1.33	3.00
	Control	1.97	0.09	0.19	0.45	0.44	2.33
	(t-stat)	(−1.46)*	(2.73)**	(3.55)**	(1.77)**	(2.14)**	(1.54)*
	(#matched pairs)	(69)	(41)	(41)	(40)	(39)	(39)
#AudComMbrs	Fraud	2.41	0.08	−0.04	0.09	0.32	2.88
	Control	2.30	0.02	0.05	0.28	0.30	2.63
	(t-stat)	(0.40)	(−0.13)	(−0.35)	(−0.73)	(−0.62)	(0.96)
	(#matched pairs)	(63)	(36)	(40)	(40)	(39)	(39)
#AudComOutsideDir	Fraud	2.34	−0.21	−0.29	−0.13	−0.24	2.71
	Control	2.25	−0.15	−0.11	−0.02	0.00	2.55
	(t-stat)	(0.19)	(−1.44)	(−1.48)	(−0.31)	(−0.26)	(0.67)
	(#matched pairs)	(63)	(36)	(40)	(40)	(39)	(39)
#FinlExperts	Fraud	0.46	0.05	0.07	0.04	−0.01	0.49
	Control	0.80	−0.04	−0.06	−0.06	0.02	0.80
	(t-stat)	(−2.40)***	(0.00)	(0.47)	(1.64)	(−0.81)	(−1.90)**
	(#matched pairs)	(63)	(36)	(40)	(40)	(39)	(39)

(*continued on next page*)

TABLE 3 (continued)

Panel C: Other Governance Variables

Variable	Firm	Year Initial[a]	1[b]	2[b]	3[b]	4[b]	Final[a]
Big4%	Fraud	0.75	0.05	0.00	−0.01	−0.01	0.74
	Control	0.85	−0.01	0.02	0.01	0.03	0.88
	(t-stat)	(−1.82)*	(1.68)*	(0.28)	(−0.77)	(−0.94)	(−2.48)**
	(#matched pairs)	(87)	(87)	(86)	(84)	(81)	(81)
CEO=COB%	Fraud	0.94	−0.05	−0.07	−0.15	−0.15	0.79
	Control	0.80	0.01	0.02	0.02	0.05	0.85
	(t-stat)	(2.79)***	(−1.92)**	(−2.40)***	(−3.51)***	(−3.86)***	(−1.04)
	(#matched pairs)	(87)	(84)	(85)	(86)	(79)	(79)
Block%	Fraud	8.73	2.74	4.40	5.72	11.06	18.67
	Control	13.38	−2.59	−0.53	1.16	2.90	15.24
	(t-stat)	(−2.05)**	(2.39)**	(2.20)**	(1.83)*	(2.39)**	(0.99)
	(#matched pairs)	(85)	(67)	(72)	(67)	(70)	(70)
InstOwn%	Fraud	0.03	−0.003	−0.005	−0.006	−0.007	0.03
	Control	0.03	0.001	0.002	0.002	0.005	0.03
	(t-stat)	(1.35)	(−1.99)*	(−2.19)**	(−2.53)**	(−1.84)*	(−0.30)
	(#matched pairs)	(22)	(18)	(18)	(17)	(13)	(13)
InsideOwn%	Fraud	25.20	−5.68	−2.81	−1.69	−1.20	22.80
	Control	26.46	−1.72	−3.89	−0.54	−2.80	22.83
	(t-stat)	(−0.42)	(−1.49)	(0.47)	(−0.49)	(0.52)	(−0.01)
	(#matched pairs)	(86)	(66)	(68)	(66)	(68)	(68)

---

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively, based on one-tailed tests, except for *BLK%* and *INST%*, for which no directional predictions are made.

[a] Initial and Final represent the first (pre-fraud detection) and last years of the analysis period, respectively; the variables for these years are measured in levels.

[b] Represent years relative to the pre-fraud detection year (Initial); the variables for these years are measured as changes from Initial.

OutsideDir% = outside director percentage;
#OutsideDir = number of outside directors;
#Directors = number of directors;
#AudComMeet = number of audit committee meetings;
#AudComMbrs = number of audit committee members;
#AudComOutsideDir = number of outside directors on audit committee;
#FinlExperts = number of financial experts on audit committee;
Big4% = proportion of firms with Big 4 audit firms;
CEO=COB = proportion of firms with the combined CEO/COB position;
Block% = percentage of shares held by > 5 percent blockholders;
InstOwn% = the percentage of shares held by institutions that file SEC Form 13F; and
InsideOwn% = the percentage of shares held by management and directors.

percentage of outside directors by more than control firms after fraud detection.[7] The change in *#OutsideDir* is also generally greater for fraud firms than for control firms over the analysis period, although this difference is statistically significant only after year 2. Over the analysis period, the change for *#Directors* is generally not different for fraud and control firms, except for year 1, where fraud firms have a larger decline than their corresponding control firms (t = –2.41). These results suggest that fraud firms improve their reputational capital through the addition of outside directors, both in number and proportion, although the decline in the total number of directors accounts for part of this result.[8]

As reported in Table 3, Panel B, I find that the mean *#AudComMeet* for fraud firms in Initial is 1.61, which is statistically smaller than the mean of 1.97 for control firms (t = –1.46). This result contrasts with that of Beasley (1996) and likely obtains due to more power in my tests—69 matched pairs in the current study versus 26 matched pairs in the Beasley (1996) study.

Table 3, Panel B also provides the cumulative numerical increase in *#AudComMeet* from its level in Initial to the fourth year afterward. The change in the mean *#AudComMeet* is greater for fraud firms than for control firms over the entire analysis period, thus providing support for the hypothesis that fraud firms increase the number of audit committee meetings by more than control firms. The mean *#AudComMeet* at Final for fraud firms is 3.00 versus 2.33 for control firms. This difference is statistically significant at the 10 percent level (one-tailed). Overall, fraud firms experience an increase of some 80 percent in *#AudComMeet* from the pre-fraud detection year to the fourth year afterward. It appears that in the wake of fraud detection the audit committee has at least visibly demonstrated that it is paying more attention to the financial reporting process.

Results provided in Table 3, Panel B also indicate that fraud firms have significantly fewer *#FinlExperts* in Initial, with 0.46 financial experts for fraud firms versus 0.80 financial experts for control firms (t = –2.40). This difference, however, persists at the end of the analysis period. I also find that while fraud firms have a majority of *#AudComOutsideDir* in the pre-fraud detection year, there is no difference between fraud and control firms in *#AudComOutside* in Initial or in any year through the fourth year afterward. One reasonable interpretation of these results is that although fraud firms' audit committees contained a majority of outsiders and some level of financial expertise, it appears that they did not pay enough attention to the financial reporting process, as evidenced by the fewer number of meetings held in Initial. Therefore, the number of audit committee meetings appears to be relatively more important than the composition of the audit committee for explaining the likelihood of committing fraud.

Results in Table 3, Panel C indicate that in the year prior to fraud detection 75 percent of fraud firms retain a Big 4 auditing firms compared to 85 percent for control firms, a statistically smaller percentage at the ten percent level (t = –1.82).[9] This result contrasts with those in DSS and Beneish (1997) and likely obtains because the severity of the frauds examined in this study likely exceeds the severity of the frauds examined in their studies in terms of the SEC rule violated. Untabulated results indicate that at the third year after fraud detection the difference between fraud and control firms relative to their use of Big 4 audit firms persists, with 75 percent of fraud firms retaining Big 4 firms versus 88 percent of the control firms, a statistically significant difference ($p = 0.03$). Overall, fraud firms are more likely than control firms to switch audit firms after fraud detection, but not audit-firm types. In the year of fraud detection, 25 percent of fraud firms switch audit firms, compared to 9 percent of control firms, a statistically significant result ($p = 0.00$). In the year after fraud detection, 24 percent of fraud firms switch audit firms, compared to 7 percent of control firms, again statistically significant ($p = 0.00$). For fraud firms, the majority of switches occurs within Big 4 firms, with 55 percent and 62 percent of the switches in the fraud detection year and year afterward, respectively. These results imply that in the post-fraud detection period fraud firms were at least able to maintain a high level of audit-firm quality, as measured by Big 4 audit firms.

---

[7] Results are qualitatively similar for tests of all governance variables holding constant the sample size at Final.

[8] The possibility exists that the results obtained for changes in outside director percentage are due to mean reversion. However, changes in outside director percentage require some explicit action by management and therefore are not like other variables, such as returns, where changes can occur without actions directed by management. I thank a referee for pointing this out to me.

[9] This result suggests that there is relatively little room for monitoring improvement along this dimension for most fraud firms. Therefore, I exclude audit-firm changes from subsequent market-based analyses.

Panel C of Table 3 also indicates that in Initial, 94 percent of fraud firms have *CEO=COB%* compared to 80 percent for control firms, a statistically greater proportion ($t = 2.79$). This result is consistent with that in DSS, who find that 86 percent of fraud firms have the combined position compared to 74 percent of control firms in the year prior to fraud detection. I find that by the end of the analysis period fraud and control firms have an indistinguishable difference in this variable. Panel C also reports that fraud firms have a significantly lower *Block%* in Initial ($t = -2.05$) compared to control firms, consistent with the finding in DSS. This difference disappears by the end of the analysis period, indicating that blockholders increase their holdings in fraud firms. I also find that while fraud firms and their matched control firms do not differ in *InstOwn%* in Initial, fraud firms experience a greater decrease in *InstOwn%* for each year of the analysis period, but do not differ in *InstOwn%* at the end of the analysis period. The sample size for this variable is very small and one therefore needs to carefully interpret these results. Finally, I find no difference in *InsideOwn%* between fraud and control firms in Initial or in any of the four years afterward.

In sum, I find that, in the year prior to fraud detection, fraud firms exhibit weakness along the governance dimensions of outside director percentage, the number of outside directors, the number of audit committee meetings, the number of financial experts on their audit committees, the quality of the external audit firm, the proportion of firms with the combined CEO/COB position, and the percentage of blockholder ownership. More interestingly, I find that by the end of the three years following fraud detection, fraud firms have indistinguishable differences from their matched control firms in outside director percentage and the proportion with the combined CEO/COB position; they also experience a *greater* number of audit committee meetings than their matched control firms. Overall, this evidence is consistent with my hypotheses and likely indicates that fraud firms wish to report their financial information more credibly. Whether this results in the restoration of fraud firms' reputations with informed capital market participants and the stock market are issues that I explore next.

Responses to Governance Changes

Overall, the results presented thus far in this paper indicate that fraud firms experience an increase of less than one outside director and an increase of one audit committee meeting in the three years after fraud detection. It is not clear how to interpret the economic significance of these results. These increases likely have a second-order effect on the restoration of the firms' reputations, whereas improving the underlying economics of these firms likely has a first-order effect. In this section, I attempt to sort out these effects and examine whether there is a relation between governance improvements and the responses of informed capital market participants and the stock market.

Informed Capital Market Participants

In this subsection, I investigate whether improvements in (1) outside director percentage, (2) the number of outside directors, and (3) audit committee activity affect informed capital market participants, such as analysts, institutional investors, and short sellers. Specifically, I test for a relation between governance improvements and whether: (1) analyst following and institutional ownership increase and (2) short sales decrease.

I obtain data on analyst following from I/B/E/S and data on institutional holdings from the CDA Investment Technologies Spectrum database, which is derived from the SEC Form 13-F disclosure forms reported quarterly to the SEC. I hand-collect monthly short interest data from the Standard & Poor's Daily Stock Price Record, except as noted for NASDAQ firms for the period subsequent to 1987. For NASDAQ firms, I obtain short interest data for the years subsequent to 1987 directly from the NASDAQ.

It is possible that the level of governance improvements affects informed capital market participants. I therefore divide fraud firms into two groups, one for changes greater than 0 in outside director percentage, the number of outside directors, and the number of audit committee meetings from the pre-fraud detection year to the fourth year afterward; and one for those changes less than or equal to 0 over the same period. I then compare the means of the associated changes in the levels of analyst following, institutional ownership, and short sales for the two groups. Results in Table 4 indicate that changes in each of the governance variables do not affect analyst following, institutional ownership, or short sales.

One possible explanation for the result with respect to institutional investors is that the governance improvements are only a necessary condition for attracting large investors and, given the severity of the frauds and the related fall-out, these investors are likely to be precluded from investing in these firms because of their fiduciary responsibilities to only invest in sound firms. It is likely, therefore, that the lingering effects of the fraud have kept institutional investors from increasing their ownership in the fraud firms. It logically follows that the lack of institutional interest likely affects analysts' decisions regarding coverage of these firms. One plausible explanation for the lack of result with respect to short interests is that these positions are likely to close out when the stock hits bottom, which is likely be around the time of fraud detection. I next investigate whether governance improvements are related to changes in the market values of the fraud firms.

Stock Returns

Prior research finds negative abnormal returns in a three-day window surrounding the announcement date of fraud (e.g., Feroz et al. 1991; DSS). If governance improvements are viewed as credible signals, then they ought to positively affect firm value through a decrease in the cost of equity capital. Given the difficulties (e.g., missing forecasts) of *directly* estimating changes in cost of equity capital associated with governance improvements, I take an indirect approach, similar in spirit to that used in DSS, and proxy for it using buy-and-hold abnormal returns.[10] That is, increases in returns would be consistent with a decline in the cost of equity capital because returns directly affect the cost of equity capital.

The results in Table 3 indicate that, for fraud firms, the number of outside directors, outside director percentage, and the number of audit committee meetings are weak relative to their corresponding control firms in the pre-fraud detection year, but strengthen by the

---

[10] In subsequent analyses, I use ROA as a proxy for expected cash flows because ROA could be driving the returns. I also test for a relation between changes in governance and corresponding changes in bid-ask spread (arguably a proxy for cost of equity capital), but fail to find one.

TABLE 4
Association of Mean Changes for Fraud Firms in Outside Director Percentage, Number of Outside Directors, and Number of Audit Committee Meetings with Mean Changes in Institutional Ownership, Short Sales, and Analyst following from the Year Prior to Fraud Detection to the Third Year Afterward

Panel A: Analysis Based on Mean in Outside Director Percentage

	Change > 0 (n)	Change ≤ 0 (n)	t-stat for Difference (p-value)
Outside Director Percentage	22.20	−14.83	9.10
	(55)	(27)	(< 0.001)
Institutional Ownership	−0.01	−0.01	0.47
	(20)	(15)	(0.64)
Short Sales	−0.01	0.00	−0.97
	(22)	(9)	(0.34)
Number of Analysts	−3.09	−1.81	−0.98
	(55)	(27)	(0.33)

Panel B: Analysis Based on Mean Changes in the Number of Outside Directors

	Change > 0 (n)	Change ≤ 0 (n)	t-stat for Difference (p-value)
Number of Outside Directors	2.22	−0.93	8.97
	(37)	(45)	(< 0.001)
Institutional Ownership	0.00	−0.01	0.21
	(16)	(19)	(0.83)
Short Sales	−0.01	−0.01	−0.24
	(11)	(13)	(0.81)
Number of Analysts	−2.68	−2.61	−0.05
	(36)	(46)	(0.96)

Panel C: Analysis Based on Mean Changes in the Number of Audit Committee Meetings

	Change > 0 (n)	Change ≤ 0 (n)	t-stat for Difference (p-value)
Number of Audit Committee Meetings	2.52	−0.47	5.37
	(34)	(18)	(< 0.001)
Institutional Ownership	−0.01	0.00	−0.87
	(16)	(7)	(0.39)
Short Sales	−0.01	0.00	−0.32
	(16)	(8)	(0.75)
Number of Analysts	−2.00	−3.84	1.18
	(34)	(18)	(0.24)

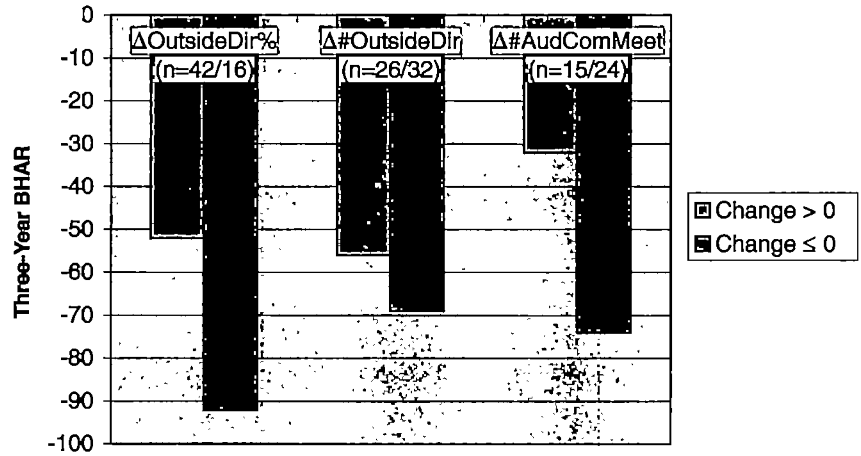
end of the fourth year afterward.[11] To investigate whether these improvements are associated with future returns, I first compare the magnitude of changes in these variables with subsequent changes in buy-and-hold returns (*BHAR*) using the CRSP equal-weighted index for the three-year period subsequent to fraud detection (see below for details on calculation of *BHAR*). I obtain returns data from CRSP. Figure 2 provides a histogram that indicates that fraud firms that make positive changes in outside director percentage, the number of outside directors, and the number of audit committee meetings have significantly ($p = 0.05$) less

---

[11] While the percentage of firms with the combined position of CEO/COB is also initially weak, only 11 out of 58 firms combined the positions of CEO and COB after the fraud detection year and 47 of the 58 firms experienced no change. Therefore, I exclude this variable from further analysis.

FIGURE 2

Buy-and-Hold Abnormal Returns for the Three Years after Fraud Detection and the Associated Governance Changes for Fraud Firms from the Year Prior to Fraud Detection to the Third Year Afterward for Frauds Detected from 1982 to 2000



Three-Year BHAR = the mean three-year buy-and-hold abnormal return for the sample of fraud firms for the three years after fraud detection using the CRSP equal-weighted index;
ΔOutsideDir% = change in outside director percentage from the year prior to fraud detection to the third year afterward;
Δ#OutsideDir = change in number of outside directors from the year prior to fraud detection to the third year afterward;
Δ#AudComMeet = change in number of audit committee meetings from the year prior to fraud detection to the third year afterward;
Change > 0 = positive governance changes from the year prior to fraud detection to the third year afterward; and
Change ≤ 0 = nonpositive governance changes from the year prior to fraud detection to the third year afterward.

negative *BHAR* than do firms that make non-positive changes.[12] Because the percentage of outside directors and the number of outside directors are likely to be highly correlated and because the effect of increasing the percentage of outside directors has a larger impact on future returns than does increasing the number of outside directors, I exclude the change in the number of outside directors from the regression analysis below.

I analyze more rigorously the relation of improvements in the quality of the board of directors and audit committee activity with abnormal returns over the three-year period following fraud detection by estimating the following regression:

$$BHAR_{i,t} = \alpha + \beta_1 \Delta OutsideDir\%_{i,t} + \beta_2 \Delta\#AudComMeet_{i,t} + \beta_3 \Delta ROA_{i,t} + \beta_4 BMV_{i,t} + \beta_5 MVE_{i,t} + e_{i,t}, \qquad (1)$$

where:

---

[12] I also use ROA as a measure of future performance, but fail to find a significant relation between changes in the governance variables mentioned above and ROA (results not presented).

$BHAR_{i,t}$ ≡ buy-and-hold abnormal return for firm *i* at time *t*, where *t* is as defined below;

$\Delta OutsideDir\%_{i,t}$ ≡ lagged change in outside director percentage from the pre-fraud detection year to the fourth year afterward for firm *i* at time *t*, where *t* is as defined below;

$\Delta \#AudComMeet_{i,t}$ ≡ lagged change in the number of audit committee meetings from the pre-fraud detection year to the fourth year afterward for firm *i* at time *t*, where *t* is as defined below;

$\Delta ROA_{i,t}$ ≡ change in net income scaled by total assets over the three years following fraud detection;

$BMV_{i,t}$ ≡ book value per share divided by the market value per share for firm *i* at time *t*, where *t* is as defined below. Book and market value are measured as of the end of the year prior to the analysis year (Source: Compustat-book value, and CRSP-market value);

$MVE_{i,t}$ ≡ log of the market value of equity for firm *i* at time *t*, where *t* is as defined below. Market value is measured as of the end of the year prior to the analysis year (Source: CRSP); and

$e_{i,t}$ ≡ the residual for firm *i* at time *t*, where *t* is as defined below.

I compute long-run buy-and-hold abnormal returns (BHAR) for fraud firm i and control firm/portfolio m as follows:

$$R_{i,t} = \prod_{t=2}^{4} (1 + r_{i,t}) - 1 \tag{2}$$

$$R_{m,t} = \prod_{t=2}^{4} (1 + r_{m,t}) - 1 \tag{3}$$

where $r_{i,t}$ and $r_{m,t}$ are the annual returns of fraud firm *i* and control firm/index *m*, respectively, on date *t*. The *BHAR* for each fraud firm is computed as the difference between the buy-and-hold returns of the fraud firm and its respective control firm/index:

$$BHAR_{i,t} = R_{i,t} - R_{m,t} \tag{4}$$

The *BHAR*s are computed with respect to (a) control firms, (b) CRSP equal-weighted index, (c) CRSP value-weighted index, and (d) CRSP S&P 500 Index. Additionally, in order to compute *BHAR* for the three years after fraud detection, only those firms with CRSP data for this three-year period are included in the following analyses.

Panel A of Table 5 reports descriptive statistics for these variables and for raw returns for fraud and control firms and the CRSP equal-weighted index for the three years following fraud detection. This panel indicates that fraud firms experience a mean (median) raw return of −22 percent (−32 percent), compared to mean (median) raw returns for control firms and the CRSP equal-weighted index of 39 percent (0.00 percent) and 31 percent (26 percent), respectively. These findings illustrate that while fraud firms experience negative abnormal returns over the analysis period, they also experience negative raw returns. This result is consistent with that provided in Figure 2.

TABLE 5

Descriptive Statistics and Pearson Pairwise Correlations for Returns, Change in Net Income, Governance Changes, and Control Variables for Fraud Firms for the Three-Year Period following Fraud Detection

Panel A: Descriptive Statistics of Returns, Governance Variables, and Control Variables (n = 34)

Variable	Mean	Std. Dev.	Median	First Quartile	Third Quartile
BHAR	−0.53	0.70	−0.83	−1.10	0.02
Fraud_Return	−0.22	0.69	−0.32	−0.71	0.52
Control_Return	0.39	1.35	0.00	−0.43	0.56
EW_CRSP	0.31	0.27	0.26	0.12	0.53
ΔROA	−0.14	0.60	0.02	−0.08	0.12
ΔOutsideDir%	6.86	12.74	5.35	0.00	12.94
BMV	8.63	50.63	0.02	0.01	0.08
MVE	4.40	1.76	4.12	3.40	5.34

Panel B: Pearson Pairwise Correlations for Returns, Change in Net Income, Governance Changes, and Control Variables (n = 34)

	BHAR	*ΔROA*	*ΔOutsideDir%*	*Δ#AudComMeet*	*BMV*	*MVE*	*Fraud_Return*
BHAR	100%	14.6%	22.6%***	7.8%	21.9%***	7.9%	92.5***
ΔROA		100%	22.5%	6.1%	13.8%	7.3%	8.4%
ΔOutsideDir%			100%	7.7%	−0.9%	−6.8%	31.8%*
Δ#AudComMeet				100%	28.2%***	9.5%	−12.7%
BMV					100%	−44.9%***	15.8%
MVE						100%	13.5%
Fraud_Return							100%

---

*,*** Significant at the 10 percent and 1 percent levels, respectively, based on two-tailed tests.

BHAR = the buy-and-hold abnormal return for fraud firms for the three years following fraud detection using the CRSP equal-weighted index;

Fraud_Return = the raw buy-and-hold return for fraud firms for the three years following fraud detection;

Control_Return = the raw buy-and-hold return for control firms for the three years following fraud detection;

EW_CRSP = the CRSP equal-weighted index return for the three years following fraud detection;

ΔROA = the change in net income scaled by total assets over the three years following fraud detection;

ΔOutsideDir% = the lagged change in outside director percentage from the pre-fraud detection year to the fourth year afterward;

Δ#AudComMeet = the lagged change in the number of audit committee meetings from the pre-fraud detection year to the fourth year afterward;

BMV = the book value per share divided by the market value per share; and

MVE = the log of the market value of equity.

Panel B of Table 5 provides the Pearson pairwise correlations for the regression variables and the raw returns for fraud firms (*Fraud_Return*) over the three-year period following fraud detection. For my sample of fraud firms, *BHAR* is positively correlated at the one percent level with *ΔOutsideDir%* (22.6 percent), *BMV* (21.9 percent), and *Fraud_Return* (92.5 percent). *ΔOutsideDir%* is also positively correlated with *Fraud_Return* (31.8 percent) at the 10 percent level. Also, *Δ#AudComMeet* is positively correlated with *BMV* (28.2 percent) at the 1 percent level. However, *ΔROA*, *Δ#AudComMeet*, and *MVE* are not significantly correlated with *BHAR* over the analysis period. In short, these pairwise correlations appear reasonable and do not seem large enough to create multicollinearity problems in the regressions presented in Table 6.[13]

Table 6 presents the results of the regressions of buy-and-hold abnormal returns on the explanatory variables in Equation (1).[14] The dependent variable is the buy-and-hold abnormal return for fraud firms for the three-year period following fraud detection based on the CRSP equal-weighted index. I obtain qualitatively similar results estimating these regressions using *BHAR* based on the CRSP value-weighted and CRSP S&P 500 indices, and control firm approach (results not tabulated). I use lagged governance variables because most of the changes in the governance variables take place by the end of the second year after fraud detection (see Table 3). Results in Table 6 confirm the relative importance of improvements in outside director percentage in explaining abnormal returns. In Model 1, the coefficient on Δ*OutsideDir%* is positive and statistically significant at the 5 percent level (t = 2.32), whereas the coefficient on Δ*#AudComMeet* is negative, but not statistically significant. This finding provides evidence that the market values increases in board independence, and is consistent in spirit with the results in Rosenstein and Wyatt (1990), but is ten times the economic significance of that found in their study. The coefficient on Δ*OutsideDir%* is 0.022, which implies that a 1 percent increase in outside director percentage is associated with a 2.2 percent increase in *BHAR*, an economically significant result. The coefficients on *BMV* and *MVE* are not statistically significant. Another interesting finding from Model 1 is that the adjusted R^2 is 18.50 percent. This very high R^2 is possibly due to Δ*OutsideDir%* being a substitute for changes in earnings.

Therefore, in Model 2, I test for the significance of Δ*ROA* in explaining returns. Results for Model 2 indicate that the coefficient on Δ*ROA* is not statistically significant, with an adjusted R^2 of 0.00 percent. In Model 3, I test for the incremental explanatory value of Δ*OutsideDir%* over Δ*ROA*. The coefficient on Δ*OutsideDir%* is again statistically significant, but at a slightly lower level of significance (t = 2.2, 5 percent level). Results also indicate that while the coefficient on Δ*ROA* is positive, it is again statistically insignificant. The overall explanatory power of Model 3 is slightly lower than that of Model 1, with an adjusted R^2 of 15.25 percent. In sum, the results from this section are consistent with the notion that improvements in outside director percentage following fraud detection are strongly linked to returns over the same period. Moreover, changes in net income over the same period provide only marginal explanatory value with respect to returns.

It is possible that the results obtained for Models 1 and 3 are driven by the correlated omitted variable of survivorship bias. That is, the results are likely to be overstated if poorly performing firms are more likely to drop out of the analysis than are those firms that are performing well (or less poorly), and if those poor performers are in the group with the smallest governance change. To test this possibility, I perform a binomial proportions test to determine whether there is a differential survivorship rate for firms in the group with the largest increase in the percentage of outside director representation and for those firms in the group with the smallest such increase. I partition the fraud sample into three equal groups based on the frequency distribution of changes in outside director percentage from

---

[13] I also conduct more formal tests of collinearity on the regressions presented in Table 6. The highest condition index for any variable is 7.8 (index values above 30 would indicate a collinearity problem).

[14] I also estimate regressions that include other controls for firm performance—lagged Δ*ROA*, Δ*EPS*, and lagged Δ*EPS*—with no qualitative difference in the results. I also include variables that interact a yearly dummy with Δ*OutsideDir%* to assess whether the speed with which governance changes were made had a differential impact on returns. The coefficients on the interaction terms are zero, with no difference among any of the years (results not tabulated).

TABLE 6

Regressions of Fraud Firms' Long-Run Buy-and-Hold Abnormal Returns for the Three Years after Fraud Detection on Governance Changes from the Year Prior to Fraud Detection to the Third Year Afterward

	Dependent Variable: *BHAR* (t-statistic)		
Variable	**(1)**	**(2)**	**(3)**
Intercept	−1.03 (−2.57)**	−1.02 (−2.26)**	−1.01 (−2.38)**
ΔOutsideDir%	0.022 (2.32)**		0.022 (2.19)**
Δ#AudComMeet	−0.05 (−1.47)		−0.05 (−1.46)
ΔROA		0.11 (0.47)	0.04 (0.20)
BMV	0.86 (0.68)	1.83 (1.35)	0.83 (0.64)
MVE	0.10 (1.33)	0.09 (1.08)	0.10 (1.26)
n	34	34	34
Adj. R^2	18.50%	0.00%	15.25%

** Significant at the 5 percent level, based on two-tailed tests.

BHAR = the buy-and-hold abnormal return for the three years following fraud detection using the CRSP equal-weighted index;
ΔOutsideDir% = the lagged change in outside director percentage from the pre fraud detection year to the fourth year afterward;
Δ#AudComMeet = the lagged change in the number of audit committee meetings from the pre-fraud detection year to the fourth year afterward;
ΔROA = the change in net income scaled by total assets over the three years following fraud detection;
BMV = the book value per share divided by the market value per share; and
MVE = the log of the market value of equity.

the pre-fraud detection year to the fourth year afterward. Results (untabulated) indicate that 32.14 percent of fraud firms with the largest increase in outside director representation drop out of the analysis versus 26.92 percent of those firms with the smallest such increase. This difference is not statistically significant ($Z = 0.42$, $p = 0.67$). Additionally, I test whether the change in outside director percentage and abnormal returns are statistically different for surviving firms and those that dropped out of the analysis. Results (untabulated) of a standard t-test indicate that there is no difference in either the increase in outside director percentage or abnormal returns for these firms ($t = -0.50$ $p = 0.62$ and $t = 0.30$ $p = 0.76$, respectively).

The results from the regressions need to be interpreted with caution. Analyses based on long-run buy-and-hold abnormal returns are fraught with well-established econometric and interpretation issues.[15] However, this paper does provide some evidence consistent with the notion that increasing outside director percentage subsequent to fraud detection yields economic benefits in the form of higher abnormal returns.

---

[15] I considered conducting a short-window event study to examine the market's response to outside director appointments, but because only a scant number of director appointments for my sample firms are disclosed prior to the issuance of proxy statements, I am unable to identify a sufficiently large enough sample to execute such a study.

V. SUMMARY AND CONCLUSIONS

This is the first study of its kind that provides evidence on the link between the credibility of the financial reporting system and the quality of governance mechanisms by investigating changes in these mechanisms subsequent to fraud detection and the corresponding economic consequences of such changes. It uses a sample of 87 firms identified by the SEC as fraudulently manipulating their financial statements. Consistent with prior research, results indicate that fraud firms have poor governance relative to a control sample in the year prior to fraud detection. Specifically, fraud firms have fewer numbers and percentages of outside board members, fewer audit committee meetings, fewer financial experts on the audit committee, a smaller percentage of Big 4 auditing firms, and a higher percentage of CEOs who are also chairmen of the board of directors. However, the results indicate that fraud firms take actions to improve their governance and that three years after fraud detection these firms have governance characteristics similar to the control firms in terms of the numbers and percentages of outside members on the board, but exceed the control firms in the number of audit committee meetings. This study also investigates whether improved governance influences informed capital market participants. The results indicate that analyst following and institutional holdings do not increase in fraud firms, suggesting that credibility was still a problem for these firms. However, the results also indicate that firms that take actions to improve governance have superior stock price performance, even after controlling for earnings performance.

This study makes several contributions to the accounting literature. It adds to our knowledge of the link between the credibility of the financial reporting system and quality of governance mechanisms by extending research on the causes and consequences of financial reporting failures (Beasley 1996; DSS; Agrawal et al 1999; Beneish 1999). This study also sheds light on the efficacy of a post-fraud detection recovery strategy that involves governance improvements by providing evidence that there are positive economic benefits that accrue to fraud firms that improve the independence of their boards. Finally, this study provides insights into the agency problem by utilizing a unique setting that provides a rare opportunity to observe agency costs directly and to study the governance measures that firms implement in an attempt to reduce these costs.

My research design, which is intended to include firms most likely to exhibit an increase in governance quality following the detection of financial reporting fraud, necessarily focuses on the most egregious GAAP violators. An issue to be resolved with future research is whether the inferences of this study are generalizable to samples of firms with a broader class of financial reporting credibility problems (e.g., earnings restatements).

REFERENCES

Agrawal, A., J. F. Jaffe, and J. M. Karpoff. 1999. Management turnover and governance changes following the revelation of fraud. *Journal of Law and Economics* 42 (April): 309–342.

Beasley, M. S. 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review* 71 (October): 443–465.

Beneish, M. D. 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16: 271–309.

———. 1999. Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74 (October): 425–457.

Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (BRC). 1999. *Report and Recommendations of The Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees.* New York, NY: New York Stock Exchange and National Association of Securities Dealers. Available at: http://www.nyse.com/pdfs/blueribb.pdf.

Bonner, S., Z–V. Palmrose, and S. Young. 1998. Fraud type and auditor litigation: An analysis of SEC accounting and enforcement releases. *The Accounting Review* 73 (October): 503–532.

Brickley, J., and C. James. 1987. The takeover market, corporate board composition, and ownership structure: The case of banking. *The Journal of Law and Economics* 30 (April): 161–180.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3: 183–199.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–36.

DeFond, M. 1992. The association between changes in client firm agency costs and auditor switching. *Auditing: A Journal of Practice & Theory* 11 (Spring): 16–31.

Fama, E. F. 1980. Agency problems and the theory of the firm. *Journal of Political Economy* 88: 288–308.

———, and M. C. Jensen. 1983. Agency problems and residual claims. *Journal of Law and Economics* 26: 327–349.

Feroz, E. H., K. Park, and V. S. Pastena. 1991. The financial and market effects of the SEC's Accounting and Auditing Enforcement Releases. *Journal of Accounting Research* 29 (Supplement): 107–142.

Jensen, M. 1993. The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance* 48: 831–880.

Karpoff, J. M., and J. R. Lott. 1993. The reputational penalty firms bear from committing criminal fraud. *Journal of Law and Economics* 36 (October): 757–802.

Klein, A. 2002. Economic determinants of audit committee independence. *The Accounting Review* 77 (April): 435–452.

Livingston, J. G. 1997. Management-borne costs of fraudulent and misleading financial reports. Ph.D. dissertation, University of Rochester.

NASDAQ Stock Market, Inc. 2004. *Listed Company Manual.* Section 303A. Washington, D.C.: Government Printing Office. Available at: http://www.nasdaq.com/about/RecentRules.stm#boards.

New York Stock Exchange (NYSE). 2004. *Listed Company Manual.* New York, NY: NYSE. Available at: http://www.nyse.com/listed/1022221393251.html.

Palmrose, Z–V. 1988. An analysis of auditor litigation and audit service quality. *The Accounting Review* 1 (January): 55–73.

Rosenstein, S., and J. G. Wyatt. 1990. Outside directors, board independence, and shareholder wealth. *Journal of Financial Economics* 26: 175–191.

Sommer, A., Jr. 1991. Auditing audit committees: An educational opportunity for auditors. *Accounting Horizons* (June): 91–93.

Sorkin, A. 2002. Tyco's chief says earnings will be lower than expected. *New York Times* (September 26): C1.

Srinivasan, S. 2004. Consequences of financial reporting failure for outside directors: Evidence from accounting restatements. Working paper, Harvard Business School.

Weisbach, M. S. 1988. Outside directors and CEO turnover. *Journal Of Financial Economics* 20: 431–460.

Yermack, D. 1996. Higher market valuation of companies with a small board of directors. *Journal of Financial Economics* 40: 185–212.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 563–583

An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation

Joseph G. Fisher
Indiana University

Laureen A. Maines
Indiana University

Sean A. Peffer
University of Kentucky

Geoffrey B. Sprinkle
Indiana University

ABSTRACT: Employment relationships provide fertile ground for both employee and employer opportunism. Employers worry about whether employees will devote sufficient effort to work, and employees are concerned about whether employers will compensate them appropriately. In this paper, we examine whether employer discretion over the size of the total employee compensation pool and the allocation of this pool among employees influences employee and employer opportunism. The results of our experiment indicate that firm output and employees' compensation are greater when the employer does not have discretion over total employee compensation, but does have discretion over the allocation of total compensation. We find that the employer's residual profit increases with discretion over the allocation of compensation among employees; however, we find no effect on residual profit of the employer's discretion over the total amount of employee compensation. Our results suggest that firms benefit from a compensation contract that establishes total employee compensation as a predetermined function of public, aggregate measures such as accounting income, but

---

We thank Walt Blacconiere, Dave Barrett, Daniel Beneish, Bob Bowen, Dave Greene, John Hassell, Dan Heitger, Frank Hodge, Jane Jollineau Kennedy, Jamie Pratt, Jerry Salamon, Terry Shevlin, Reed Smith, Michael Williamson, two anonymous reviewers, and participants at research workshops at Emory University, Indiana University Bloomington, and Indiana University Indianapolis, University of Maryland, Michigan State University, University of Missouri, The University of North Carolina at Chapel Hill, University of South Carolina, and University of Washington for helpful comments and suggestions. The University of Kentucky and the Pressler Faculty Fellowship at Indiana University provided generous funding for this project.

Editor's note: This paper was accepted by Marlys Gascho Lipe, Editor.

Submitted July 2003
Accepted Sept 2004

provides the employer at least some discretion to allocate this compensation using private information. However, our results caution that employees and employers may not have similar preferences for the degree of employer discretion over the determination of total employee compensation.

Keywords: *compensation; discretion; subjective performance evaluation; double-sided moral hazard.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

Employment relationships typically are characterized by two-sided concerns regarding opportunistic behavior (Demski 1997, 579). Employers are concerned about whether employees will work diligently, while employees worry about whether employers will reward this work appropriately. The potential for employees and employers to act in their own self-interest when making these choices threatens to destroy the cooperation needed for productive firm outcomes.

Opportunism may not arise if workplace norms result in cooperation among employers and employees. However, firms typically do not rely solely on workplace norms to preempt opportunism, and implement control systems, such as performance evaluation and compensation systems, to mitigate opportunism. These systems govern two compensation decisions: (1) how to allocate total remuneration between the employer and employees and (2) how to allocate total employee remuneration among employees (Baiman and Rajan 1995). The first decision establishes the *size* of the total employee compensation pool and the second determines the *allocation* of this pool among employees. The approach a firm uses to make these decisions can influence both employer and employee opportunism.

In this paper, we examine whether an employer's discretion over the two compensation decisions affects employee and employer opportunism. The effect of employer discretion over the two compensation decisions is unclear because discretion creates tension with respect to employer and employee opportunism. An employer can use discretion to her personal benefit, increasing her share of total compensation at the expense of employees or allocating compensation to "favored" employees.[1] However, discretion also allows an employer to reduce employee opportunism by determining employee compensation using measures and approaches that best reflect employees' collective and individual efforts.

The tension between employer discretion and employer/employee opportunism typically is connected to firms' information systems. Firms' formal information systems, such as accounting systems, provide information that is publicly observable or verifiable but often reflects only aggregate employee performance (Baiman and Rajan 1995). Employers can alleviate employees' concerns about employer opportunism by giving up discretion and committing *ex ante* to a total compensation pool and its allocation among employees based on such public information. However, public information may not reflect all relevant aspects of performance. In particular, aggregate information provides little basis for employers to differentially reward individual employee performance and prevent employees from free-riding on the efforts of other employees. Thus, compensation approaches based on publicly available information may reduce employer opportunism but may not prevent employee opportunism.

---

[1] Throughout the paper, we use feminine pronouns to refer to employers and masculine pronouns to refer to employees.

Firms also have informal sources of information. This information can reflect multiple aspects of aggregate performance and/or isolate the performance of individual employees, but is privately observed by the employer. An employer can reduce employees' opportunism by retaining discretion to use her private information in performance evaluation. For example, an employer can reduce the benefit of free-riding by using her private information on individual performance to allocate the compensation pool among employees. Unfortunately, allowing the employer discretion to use private information generates employee concerns that the employer will fail to reward employees appropriately for their aggregate or individual efforts. Thus, a compensation scheme in which an employer has full discretion to use private information may reduce employee opportunism but allows for employer opportunism.

The relative strengths and weaknesses of restricting versus allowing employer discretion suggests that there may be complementary roles for discretion versus nondiscretion with respect to the two compensation decisions. Specifically, an approach in which the employer does not have discretion over the size of the total employee compensation pool (hereafter pool size) but does have discretion over the allocation of the employee compensation pool (hereafter pool allocation) may alleviate concerns about both employer and employee opportunism (Baiman and Rajan 1995). This approach suggests that there also are complementary roles for public and private information—public information provides the basis for the total compensation pool and private information assists in the allocation of this pool.

To investigate the effects of employer discretion on employee and employer opportunism, we examine situations in which the employer either has full discretion or no discretion over the *size* of the employee compensation pool and/or the *allocation* of this pool among employees. We use a laboratory experiment with a 2 (pool size discretion) × 2 (pool allocation discretion) × 8 (period) mixed design. In the experiment, the employer either has discretion to determine the size of the two-employee compensation pool (discretionary pool size, DS) or the pool is set at two-thirds of the revenue from total employee output (nondiscretionary pool size, NDS). Similarly, the employer either has discretion in allocating the compensation pool between the two employees (discretionary pool allocation, DA) or the pool is split evenly between the two employees (nondiscretionary pool allocation, NDA). The employer receives the residual profit after subtracting employee compensation from total revenue.

In the experiment, 237 undergraduate business students, in groups of three, acted either as an employer or as one of two employees. Employees allocated three minutes between a work task (decoding numbers into letters) and compensated leisure for each of eight periods. After each work session, employers in the discretionary conditions determined the size of the employee compensation pool and/or the allocation of this pool between the two employees. Employers received noisy private information about individual employee output to use if desired in making their compensation decisions. Thus, in the discretionary conditions, employers had the opportunity to more closely align an employee's compensation with individual production than in the nondiscretionary conditions in which compensation was a multiple of total output and/or evenly split between the two employees.

We consider the impact of employer discretion over pool size and pool allocation on total output, as well as the division of the value of total output (i.e., revenue) between employer remuneration and employee remuneration. The results indicate that both total output and employee compensation benefit from restricting the employer's discretion over the size of the compensation pool, and allowing the employer to have discretion over the allocation of this pool. Total output and employee compensation were greater when the employer did not have discretion over pool size than when she did have discretion, as well when the employer had discretion over the pool allocation than when employees received equal compensation. These results hold even when we eliminate the last two periods to control for possible end-game effects.

Our results for the effect of discretion on the employer's residual profit are mixed. Over the entire eight periods, the employer's residual profit did not differ between the discretionary and nondiscretionary pool size conditions. The employer's residual profit was greater in the discretionary pool allocation condition than in the nondiscretionary condition. However, we found no effect of either pool size or pool allocation discretion on employer's residual profit after eliminating the final two periods. This result suggests that discretionary pool allocation primarily mitigates employee opportunism that arises as the employment relationship nears its end.

Our results also document a fairly high level of cooperative behavior between employers and employees in the absence of controls that mitigate opportunism. In the DS-NDA contract, which provides no explicit controls to mitigate employee or employer opportunism, employees produced, on average, 75 percent of the amount produced in the best-performing NDS-DA contract. Additionally, employers with discretion over pool size paid out, on average, 59 percent of revenues as compensation to the two employees. However, the effects of opportunism associated with either a discretionary pool size or nondiscretionary pool allocation approach increased over the eight periods. For example, in the last period, output in the DS-NDA contract was only 45 percent of that in the NDS-DA contract. Moreover, employers in the two discretionary size conditions paid only 51 percent of revenue to employees in the last period.

This study contributes to the control systems literature by examining implications of employer discretion in performance evaluation and compensation for employer and employee opportunism. Our results document a fairly high level of productive output in the absence of specific controls related to employer discretion/nondiscretion intended to mitigate employee and employer opportunism. Rather, employers and employees appear to recognize the economic benefits of joint cooperation and rely on social values, such as trust and reciprocity, to keep productive outcomes intact over all but the final periods of employment (Axelrod 1984; Berg et al. 1995; Sprinkle 2003). These findings highlight that the efficacy and cost of performance evaluation and compensation arrangements should be compared not simply against each other, but against the cost and efficacy of naturally occurring social mechanisms in the workplace.

Despite these cooperative outcomes, our study documents benefits of control systems with explicit incentives to mitigate employer and employee opportunism. Specifically, allowing the employer discretion to allocate compensation among employees increases firm output and both employee and employer remuneration. However, the implications of our results for discretion in setting the size of total employee compensation are less clear. While eliminating employer discretion over the pool size improves output and employee compensation, the employer's residual profit is not greater under the nondiscretionary pool size case than under the discretionary pool size case. In essence, the benefit to the employer of increased output associated with a nondiscretionary pool size is not sufficiently large to offset the cost to the employer of having to pay greater per unit compensation to employees in our experiment. Thus, employers may prefer to retain discretion over both pool size and pool allocation, to the detriment of firm output and employee compensation. However, given the compensation-related benefits to employees of a nondiscretionary approach to pool size, competition among employers for employees could lead employers to relinquish their discretion over compensation pool size.

Finally, our study complements research on the use of contractible (public) versus noncontractible (private) information in compensation contracts. Analytical research documents the benefits of compensation contracts that base the total pool size on contractible information and the pool allocation on employer's private information (Baiman and Rajan 1995). Our results generally provide empirical support for Baiman and Rajan's (1995) conclusion. Since contractible information often emanates from formal accounting systems, our findings also suggest that there are complementary roles in designing compensation systems for information from formal accounting systems and private information from a firm's informal information systems.

The remainder of this paper is organized into four sections. The next section provides background and develops the hypotheses, and Section III describes the experiment used to test these hypotheses. Section IV presents the results, and Section V provides a summary and discussion of the results.

II. BACKGROUND AND HYPOTHESES

Employer and Employee Opportunism—Discretionary Pool Size/Nondiscretionary Pool Allocation Contract (DS-NDA)

We examine a setting in which employees work independently on a production task for a finite number of periods. Each employee allocates a fixed amount of time each period between work and leisure, and derives utility from both leisure time and compensation. Output for the group of employees is the sum of individual employee output. Firm revenue equals group output multiplied by the competitive market price. The employer is a residual claimant who receives total firm revenue less the amount paid as compensation to employees.

The firm has an information system that measures employee performance and a compensation system that rewards employee performance (Zimmerman 2003, 169). To study our research question on the effects of employer discretion on opportunism, we vary characteristics of these two systems. We first examine a setting that provides no formal controls to mitigate either employer or employee opportunism and then add controls that potentially mitigate one or both types of opportunism.

In our initial setting, the information system measures group output (e.g., revenue or income for the entire company or a division), but not individual employee output.[2] The employer retains discretion to determine the size of employees' compensation pool. Given the lack of information on individual performance, the employer does not retain discretion over the allocation of the compensation pool. Rather, the pool is divided equally among the employees. We term this compensation system the discretionary size/nondiscretionary allocation (DS-NDA) contract.

These choices for information and compensation systems represent company settings that provide fertile ground for both employer and employee opportunism. Employees have an incentive to act opportunistically and shirk in their work effort because their individual compensation is based on group performance. If an employee receives more utility from leisure than from the expected incremental compensation he would receive from working, then he may shirk and receive work compensation based on the efforts of his co-workers (free-ride).[3]

---

[2] In practice, employers often observe aggregate performance measures from firms' formal information systems, but can not disentangle individual employee performance due to complexities of the production process. In these settings, employers frequently use aggregated performance as a basis for incentive contracts and allocate total compensation in a manner unrelated to individual performance (Milgrom and Roberts 1992, 413).

Employers also have an incentive to act opportunistically given their discretion over the size of total employee compensation. Given a finite-period world, a self-interested employer will pay no employee compensation in the last period since it only reduces her residual payment and has no incentive value for employees' future work-leisure allocation decisions (employees have already made all such decisions). If employees believe the employer will not pay compensation in the final period, then they have no incentive to work and will allocate all time to leisure. Backward induction implies a Nash equilibrium in which the employer does not pay compensation and employees do not work in any period.[4]

Both employer and employee opportunism threaten to destroy productive outcomes that could benefit all firm members. This threat arises from employer discretion over total employee compensation and employee discretion over work effort, combined with incentives for employers and employees to use their discretion for personal gain.

The threat of opportunism may not materialize if employees and employers choose to cooperate with one another, either out of recognition of their mutual benefit or out of fear that others will retaliate against opportunistic behavior (Fehr and Gächter 2000). Such cooperative behavior requires trust that others will not act opportunistically. Research on trust games shows that individuals are willing to place themselves at risk by cooperating and trusting that other individuals will not exhibit self-interested behavior (Berg et al. 1995). If all employees initially trust that other employees will choose to work rather than to shirk, then a productive equilibrium can be sustained as long as all employees reciprocate. Indeed, Fisher et al. (2003) find that employees' work allocations and performance are significantly greater than zero in a setting that does not provide a compensation-based mechanism to mitigate employee opportunism.[5]

Employees also must trust that the employer will not act in her own self-interest, but instead will reciprocate by compensating employees appropriately. This expectation is not without foundation, as research indicates that employers reward employees who exert high levels of effort (Fehr et al. 1997). Additionally, over time, an employer can develop a reputation for compensating employees that induces employees to work in advance of receiving compensation (Bull 1987).

Thus, it is possible that extreme opportunism will not arise in the DS-NDA contract due to social values such as trust and reciprocity. However, if these social values are not sustainable workplace norms, then other mechanisms are needed to mitigate employee and employer opportunism. The firm's performance evaluation and compensation system can serve as such a mechanism. In the next section, we use implications of a paper by Baiman and Rajan (1995) to consider how changes to the employer's discretion in setting employee compensation can help to mitigate employer and employee opportunism. Specifically, we examine three compensation contracts that change one or both aspects of the DS-NDA contract with respect to employer discretion/nondiscretion in determining the size of the employee compensation pool and/or the allocation of this pool among employees.

---

[3] This is similar to public goods settings in which individuals have the incentive not to contribute to the common pool, but consume the benefits of contributions from other individuals (Ledyard 1995).

[4] For example, the employer knows that paying compensation in the next-to-last period cannot motivate employees' effort for that period since production has already occurred, but could motivate employees' effort in the last period. However, the employer knows that employees will not work in the last period given that she can not commit to paying compensation. Thus, there is no incentive for her to pay compensation in the next-to-the-last period.

[5] However, Fisher et al. (2003) include a production goal in their setting, which may act as a separate mechanism influencing employee behavior.

Compensation Contracts to Mitigate Employer and Employee Opportunism

Baiman and Rajan (1995) provide insights on features of compensation contracts that have potential to mitigate employee and employer opportunism. Baiman and Rajan (1995) (hereafter B-R) examine a contract in which the employer commits *ex ante* to determine the size of the employee compensation pool based on jointly observable (contractible) information, but does not commit to a specific pool allocation. Rather, the employer has access to private (noncontractible) information about an individual employee's performance and retains the right to use this information to allocate the compensation pool. B-R analytically show that this contract results in a strict Pareto improvement over a compensation contract in which the employer commits to determining compensation using only jointly observable information.

The findings in B-R highlight the importance of employer discretion for employee and employer opportunism associated with the DS-NDA contract. The optimal contract in B-R has two features that differ from the DS-NDA contract. First, in the B-R contract, the employer gives up discretion to set the compensation pool size *ex post*. Second, the employer retains discretion to allocate the compensation pool among employees using her private information related to individual employee performance. Thus, B-R's analytical results suggest that employee and employer opportunism can be mitigated by choices related to employer discretion. We discuss these issues below for employee and employer opportunism.

Mitigating Employee Opportunism—Discretionary Pool Size/Discretionary Pool Allocation Contract (DS-DA)

In the DS-DA contract, the employer has complete discretion over both the amount of the compensation pool and the allocation of the pool among employees. While the DS-DA contract does not alleviate employer opportunism due to this discretion, the contract provides an opportunity for mitigating employee opportunism. Specifically, if the information system is expanded to include the employer's private noisy signals about each employee's output (as in B-R), the employer can use her discretion to determine the allocation of compensation among employees based on private information about individual employee performance.[6] By more closely aligning an employee's compensation with his individual performance, the employer reduces the benefits of shirking to the employee, which should increase his work efforts.

Mitigating Employer Opportunism—Nondiscretionary Pool Size/Nondiscretionary Pool Allocation Contract (NDS-NDA)

In this contract, the employer has no discretion in setting either the size of the compensation pool or its allocation among employees; in essence, the employer has no role in determining employee compensation. The total employee compensation pool is a function of output (percentage of total revenue) and employees receive an equal share of the compensation pool. The equal division of the compensation pool among employees provides no formal controls for employee opportunism; however, the NDS-NDA contract does mitigate employer opportunism by taking away the employer's discretion. Note that the NDS-NDA contract is more extreme in this respect than the optimal B-R contract. The employer has no discretion in the NDS-NDA contract, while the employer in the B-R contract retains discretion over the parameters of the compensation pool but commits *ex ante* to these parameters using jointly observable information. Thus, the NDS-NDA contract focuses on eliminating employer opportunism, as opposed to creating an optimal contract from the employer's viewpoint, as in B-R.[7]

---

[6] The characterization of the employer's private information as a noisy signal of individual employee performance reflects the idea that individual employee output typically is not easily identified (Prendergast 1999, 11). However, employers often receive information related to aspects of individual employee performance. This information typically is private and based on subjective assessments. If the signals were public rather than private, then an employer could explicitly tie an employee's compensation to his individual performance signal.

Mitigating Both Employee and Employer Opportunism—Nondiscretionary Pool Size/ Discretionary Pool Allocation Contract (NDS-DA)

In the NDS-DA contract, the compensation pool is a predetermined function of group output, but the employer has discretion in allocating the pool among employees. Employer opportunism is mitigated because the employer has no discretion over the size of the compensation pool. Moreover, the employer can reduce employee opportunism by using her private signals about individual performance to align an employee's compensation with his performance. Thus, the NDS-DA contract has potential to mitigate both employer and employee opportunism.

Hypotheses

The prior discussion indicates that use of a nondiscretionary approach to determining the size of the employee compensation pool can mitigate employer opportunism. This increases employees' willingness to devote time to work, which, in turn, increases output. Similarly, the use of a discretionary approach to allocating the employee compensation pool should mitigate employee shirking and increase output. Employees will devote more time to work because they expect that their employer will exercise her discretion and use information about their individual production to determine their share of total employee compensation. *Ceteris paribus*, both employees and employers should benefit from the increased output. These arguments lead to the following hypotheses.

Employees' Group Output (Performance)[8]

H1a: Group output will be greater when the employer does not have discretion over the size of the employee compensation pool than when the employer has discretion over the size of the employee compensation pool.

H1b: Group output will be greater when the employer has discretion over the allocation of the employee compensation pool than when the employer does not have discretion over the allocation of the employee compensation pool.

---

[7] There are several other differences between B-R and our paper. First, the B-R model is a single-period model while our setting includes a finitely repeated setting. Second, in the B-R model, outcomes and signals are dichotomous, while in our setting production outcomes and signals can take on numerous levels. Finally, in the B-R model, there are jointly observable signals about each employee's individual performance, whereas in our setting all individual signals are the employer's private information and only group output is jointly observable. Thus, while B-R has implications for our setting, our paper is not a direct test of their analytical predictions.

[8] Predictions similar to H1a and H1b could be made for employees' allocation of time to work. Given our assumption that output is a function of work time, we provide hypotheses only for output (performance). We report empirical results for employees' work time in footnote 13.

Employees' Compensation

H2a: Employees' compensation will be greater when the employer does not have discretion over the size of the employee compensation pool than when the employer has discretion over the size of the employee compensation pool.

H2b: Employees' compensation will be greater when the employer has discretion over the allocation of the employee compensation pool than when the employer does not have discretion over the allocation of the employee compensation pool.

Employer's Residual Profit

H3a: The employer's residual profit will be greater when the employer does not have discretion over the size of the employee compensation pool than when the employer has discretion over the size of the employee compensation pool.

H3b: The employer's residual profit will be greater when the employer has discretion over the allocation of the employee compensation pool than when the employer does not have discretion over the allocation of the employee compensation pool.

Notwithstanding these hypotheses, there are at least two behavioral forces working against the predictions. First, as noted earlier, workplace norms related to cooperation, trust, and reciprocity may naturally alleviate employer and employee opportunism, leaving little role for compensation contracts to mitigate such opportunism. Second, with respect to H1b, H2b, and H3b, the effectiveness of discretionary allocation to mitigate employee opportunism depends on whether an employer uses her private information in a manner that motivates employee effort (B-R). While ideally the employer will align an employee's compensation with his individual performance, research suggests that employers are biased in performance evaluation (Prendergast 1999). For example, employers may be unwilling to make significant distinctions between employees' contributions and compress the individual signals, leading to a centrality bias. Alternatively, employers may make more extreme distinctions between employees than is evident in their performance signals, resulting in a halo effect for higher-performing employees.

III. METHOD

Participants and Design

Two hundred thirty-seven undergraduate students participated in our computer-based laboratory experiment, which employed a 2 (pool *size* discretion) × 2 (pool *allocation* discretion) × 8 (period) mixed design.[9] We manipulated pool size discretion and pool allocation discretion between-subjects and periods (work sessions) within-subjects. We randomly assigned participants to one of the four compensation contract conditions: discretionary pool size/nondiscretionary pool allocation, discretionary pool size/discretionary pool allocation, nondiscretionary pool size/nondiscretionary pool allocation, and nondiscretionary pool size/discretionary pool allocation.

---

[9] Two hundred forty-three students participated in the experiment; we eliminated two groups (six participants) due to computer-related malfunctions that resulted in missing data for some members of these groups.

Procedure

One-third (two-thirds) of the participants assumed the role of employer (employee) in the experimental task. For each of the four compensation contracts, the administrator randomly matched one employer with two anonymous employees. The first few computer screens explained the production task, which involved decoding numbers into letters (Chow 1983). Participants then received a description of one of the four compensation schemes tailored to their experimental condition, along with several numerical examples. Each employee's compensation contract took the following generic form in each period:

$$TP_i = [a \times (T_i - W_i)] + B_i \tag{1}$$

where:

TP_i = total pay for employee i, $i = 1, 2$;
a = pay per unit of leisure time consumed;
T_i = total time available for employee i;
W_i = time allocated to work by employee i; and
B_i = work compensation for employee i.

The employer's residual profit in each period was computed according to the following generic formula:

$$RP = [r \times (O_1 + O_2)] - (B_1 + B_2) \tag{2}$$

where:

RP = residual profit received by the employer;
r = revenue per unit of output;
$O_1 + O_2$ = output of the two employees (group output); and
$B_1 + B_2$ = total work compensation for the two employees.

In each of the four contracts, $T_i = 180$ seconds, $a = 0.5$ francs, and $r = 3.54$ francs (one franc, our experimental currency, equaled one cent). For the two nondiscretionary pool size contracts (NDS-NDA and NDS-DA), the total employee compensation piece rate was 2.36 francs per unit of group output.

Our choice of parameter values was driven by the determination of a piece rate that provides incentives for employees to shirk in the NDS-NDA contract. For employees to want to shirk, the leisure pay per second must exceed the compensation an employee would receive for devoting an additional second to work. We fixed leisure pay at $a = 0.5$ francs per second. We then determined a piece rate for each employee of 1.18 francs per unit based on: (1) pilot-testing in which participants decoded 58 numbers on average in three minutes (approximately 0.32 numbers per second) and (2) a marginal per capita return (MPCR) of 0.75 based on a common value used in public goods games (Ledyard 1995). Thus, an average participant in the NDS-NDA contract could earn either 0.5 francs from allocating one second to leisure or 0.38 (1.18 × 0.32) francs from allocating one second to work, leading to an incentive to shirk.[10]

The contract description varied for each condition as follows.

- In the DS-NDA compensation contract, B_i was described as the total compensation pool determined by the employer divided by two (i.e., split equally between the two employees).
- The DS-DA compensation contract was described by Equations (1) and (2), with a note that the employer had discretion to determine B_i for each employee.
- In the NDS-NDA compensation contract, B_i was described as $1.18 \times (O_1 + O_2)$, where 1.18 francs equaled the nondiscretionary amount of individual employee compensation per unit of group output. This amount represented an equal division of revenue among the three firm members (3.54/3). Thus, the total employee compensation pool for the two employees was computed as 2.36 per unit of group output.
- In the NDS-DA compensation contract, B_i was described as $(d_i) \times (2.36) \times (O_1 + O_2)$ where d_i equaled the fraction of the total compensation pool the employer decided to give to employee i ($d_1 + d_2 = 1$) and 2.36 francs equaled the nondiscretionary amount of total employee compensation per unit of output.

The computer screens then described the task. For each of eight work sessions, each employee allocated 180 seconds between work and leisure. After every work session, each employee learned the number of items he decoded correctly. The employer received a private signal for each employee, termed reported output. Reported output equaled the actual number of items correctly decoded by the employee adjusted by a randomly chosen number between −7 and +7 using a uniform distribution, and reflects the typical situation in which employers receive noisy measures of individual employee's performance.[11] The employer and employees learned total reported group output (i.e., the sum of reported output for the two employees); however, only the employer knew reported output for each individual employee.

All participants completed three three-minute training sessions to ensure they understood the task. Additionally, prior to starting the experimental sessions, participants answered several manipulation check questions to ensure that they understood the incentive contracts, the information received by employees and employers, and the decisions to be made by employees and employers. If a participant answered a question incorrectly, then the computer returned to the screen containing the appropriate information. Participants had to answer all questions correctly before proceeding to the work sessions.

After each work period, the computer automatically determined the pool size and/or pool allocation in the nondiscretionary conditions. The employer determined the pool size and/or pool allocation in the discretionary conditions. Employees' compensation and the employer's residual profit were based on reported, rather than actual, output because participants learned total reported, but not total actual, output. After completing all eight periods, participants completed an exit questionnaire and received their compensation.

---

[10] As an example, if both employees fully consume leisure, each would earn 90 francs (180 seconds × 0.5 francs) in every work session. If two average employees allocate all time to work in a session, each would earn 137 francs (58 units × 2 persons × 1.18 francs/unit). However, if one employee engages in free riding, he could earn 90 francs from leisure plus 68 francs if the other employee works for the entire period (58 units × 1.18 francs per unit), for a total of 158 francs. Thus, free riding is the dominant strategy when total employee compensation is split evenly between the two employees, although output is maximized when both employees work.

[11] Each employee's reported output can be viewed as a function of the number of items he decoded correctly and a random state of nature. Reported output was constrained to be greater than or equal to zero.

Dependent Measures

We examine three dependent measures to test our hypotheses: employees' group performance (number of items actually decoded correctly), employees' group compensation, and employer's residual profit. We use actual, rather than reported, performance as a dependent measure because it reflects employee output without the effect of the random error. For employees' group compensation, we combine the corresponding amounts for the two employees paired with one employer (described above in Equation (1)) to reflect the group's total compensation. The employer's residual profit was calculated as shown in Equation (2).

IV. RESULTS

Presentation of Experimental Results

We present our results in three tables, one for each dependent measure: employees' group performance (actual output), employees' group compensation, and employer's residual profit.[12] Panel A of each table presents means over the eight periods for each dependent measure by pool size discretion and pool allocation discretion condition. Panel B of each table contains the related ANOVA, with pool size discretion and pool allocation discretion as the between-subjects factors and period as the within-subjects (repeated measures) factor. Additionally, Figure 1 presents graphical results for all eight periods for the three dependent measures.

Employees' Performance (Hypotheses 1a and 1b)

Table 1 and Panel A of Figure 1 present results for employees' actual group performance (output). Hypothesis 1a predicted that employee group performance would be greater in the nondiscretionary pool size conditions than in the discretionary pool size conditions. Hypothesis 1b predicted that employee group performance would be greater under discretionary than nondiscretionary pool allocation. The results in Table 1 support these predictions.

Panel A of Table 1 indicates that the mean performance over all eight periods for the nondiscretionary pool size conditions was 111.08, compared to mean performance of 95.02 for discretionary pool size conditions ($F = 9.90$, $p < 0.01$, as shown in Panel B), supporting H1a. As predicted by H1b, the mean performance of 110.43 for the discretionary pool allocation conditions was higher than 95.28 for the nondiscretionary pool allocation conditions ($F = 8.64$, $p < 0.01$).

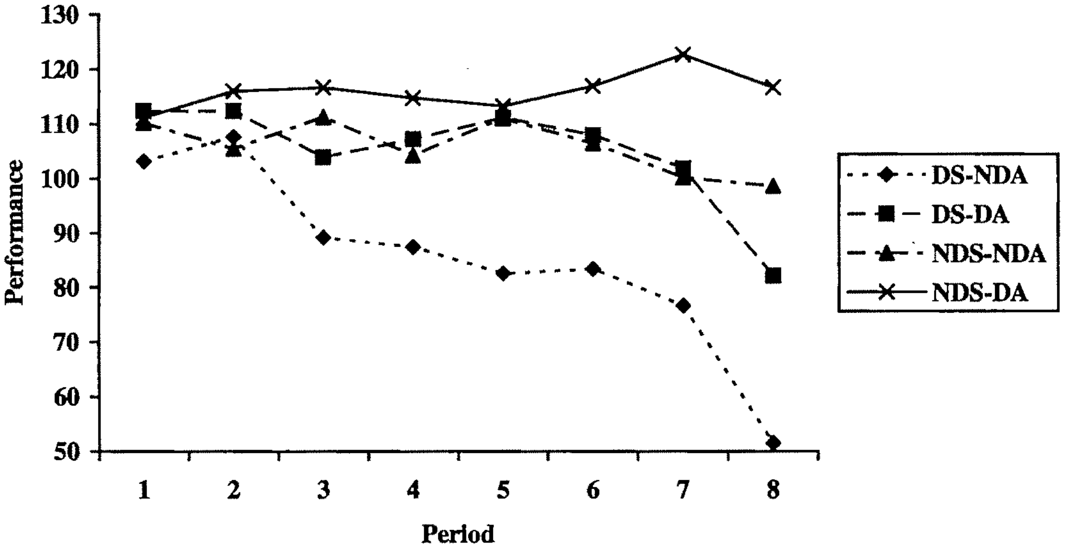
The ANOVA in Panel B of Table 1 shows a statistically significant main effect for period ($F = 10.67$, $p < 0.01$), a statistically significant period-by-pool size discretion interaction ($F = 8.04$, $p < 0.01$), and a statistically significant period-by-pool allocation discretion interaction ($F = 2.61$, $p < 0.02$). As indicated in Panel A of Figure 1, performance in the first period was roughly equal across the four contracts (between 103 and 112 units); however, by the final period, there was significant variation in performance (a low of 52 for the DS-NDA contract and a high of 117 for the NDS-DA contract). Thus, in the

---

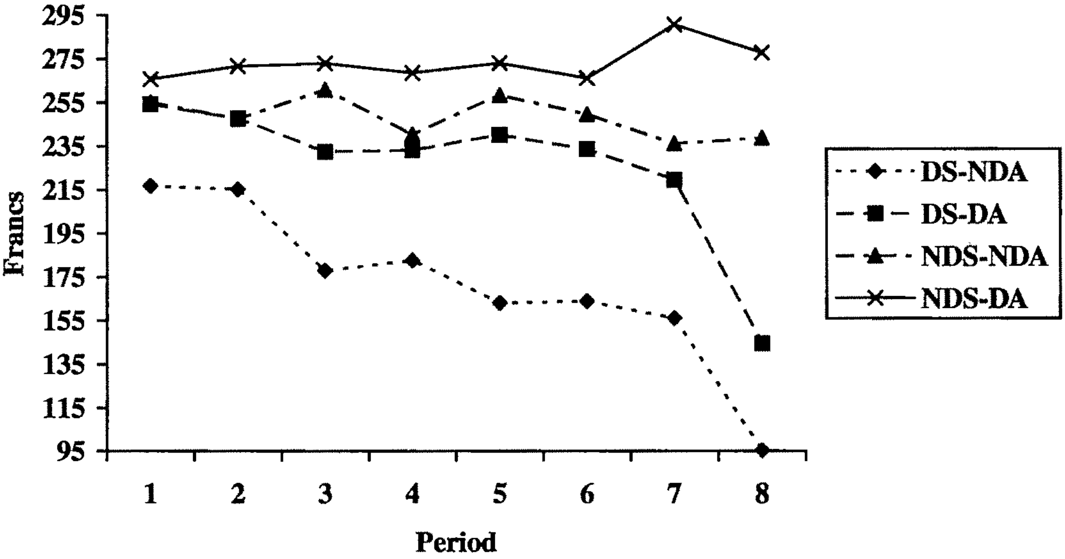
[12] Table 1 reports actual output, while Table 2 reports employee group compensation that is based on reported output (actual output adjusted by a randomly selected number between −7 and +7). For this reason, group compensation for the nondiscretionary pool size conditions (NDS-NDA and NDS-DA) in Table 2 does not equal output in Table 1 multiplied by 2.36 francs.

FIGURE 1
Effect of Compensation Contracts on Employee Group Performance, Employee Group Compensation, and Employer Residual Profit by Period[a]

Panel A: Average Employee Group Performance by Period[b]



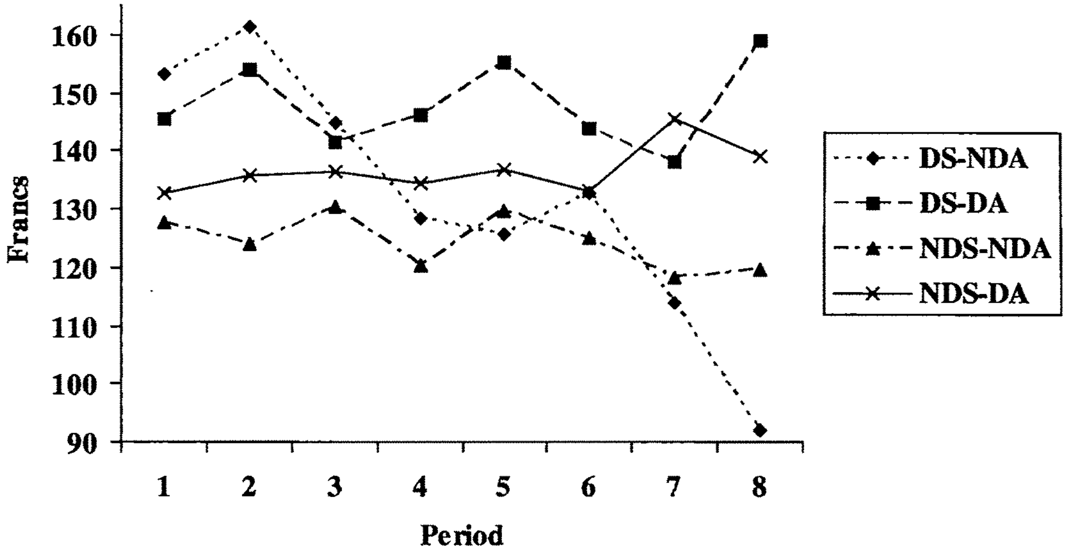
Panel B: Average Employee Group Compensation by Period[c]



(*continued on next page*)

FIGURE 1 (continued)

Panel C: Average Employer Residual Profit by Period[d]



[a] The four contracts are discretionary size/nondiscretionary allocation (DS-NDA), discretionary size/ discretionary allocation (DS-DA), nondiscretionary size/nondiscretionary allocation (NDS-NDA), and nondiscretionary size/discretionary allocation (NDS-DA).

[b] Employee group performance equals the sum across the two employees in each employer-employee group of the actual number of items correctly decoded by the employee.

[c] Employee group compensation equals the sum of the compensation received by the two employees in each employer-employee group.

[d] Employer's residual profit equals total revenue less compensation paid to the two employees in each employer-employee group.

absence of explicit incentives, the effect of employee and employer opportunism had a greater impact on performance as the employment relationship approached its end.[13]

Employees' Compensation (Hypotheses 2a and 2b)

Table 2 and Panel B of Figure 1 present results for employees' group work compensation (the amount of firm revenue received by employees).[14] Hypothesis 2a predicted that employee compensation would be greater in the nondiscretionary than discretionary pool

---

[13] An ANOVA for reported group performance showed identical inferences to those reported in Table 1 for actual group performance. Additionally, an ANOVA for employee work time indicated statistically significant main effects for pool size discretion, pool allocation discretion, period, the period-by-pool size discretion interaction, and the period-by-pool allocation discretion interaction. Similar to results for performance, opportunism in employees' time allocations became more apparent in the later periods. In the first period, employees allocated about the same percentage of time to work in the worst (DS-NDA) contract and best (NDS-DA) contract (90 percent and 94 percent, respectively). By the final period, however, employees in the DS-NDA contract allocated only 41 percent of their time to work, compared to 90 percent for the NDS-DA contract.

[14] We also performed an ANOVA with the dependent variable defined as employees' work compensation plus leisure pay. This ANOVA more completely reflects employees' total benefits, i.e., the benefits that employees receive from both their work effort and their opportunism (time allocated to leisure). Inferences from ANOVA results are the same as those reported in Panel B of Table 2 for work compensation.

TABLE 1
Employee Group Performance (H1a and H1b)

Panel A: Means (Standard Deviations) for Employee Group Performance over All Periods[a]

Compensation Pool Allocation[c]	Compensation Pool Size[b]		
	Nondiscretionary	**Discretionary**	**Overall**
Nondiscretionary	105.92	85.18	95.28
	(21.15)	(26.20)	(25.79)
	n = 19[d]	n = 20	n = 39
Discretionary	115.99	104.86	110.43
	(14.80)	(25.86)	(21.55)
	n = 20	n = 20	n = 40
Overall	111.08	95.02	102.95
	(18.64)	(27.56)	(24.79)
	n = 39	n = 40	n = 79

Panel B: ANOVA on Employee Group Performance

Source of Variation	SS	df	MS	F[e]	p[f]
Between-Subjects					
Pool Size Condition (SC)	40,102	1	40,102	9.90	< 0.01
Pool Allocation Condition (AC)	34,982	1	34,982	8.64	< 0.01
SC × AC	3,645	1	3,645	0.90	< 0.35
Group(SC × AC)	303,674	75	4,049		
Within-Subjects					
Period[g]	28,147	7	4,021	10.67	< 0.01
Period × SC	21,211	7	3,030	8.04	< 0.01
Period × AC	6,888	7	984	2.61	< 0.02
Period × SC × AC	2,937	7	420	1.11	< 0.36
Period × Group(SC × AC)	197,744	525	377		

---

[a] Employee groups consist of two employees paired with one employer. Employee actual group performance equals the sum across the two employees of the actual number of items correctly decoded by the employee.

[b] We manipulated the determination of the compensation pool size at two levels: nondiscretionarily set at 2.36 francs (cents) per unit of output or discretionarily set by the employer. In the nondiscretionary case, employees in total received two-thirds of revenue (3.54 per unit of output) and the employer received one-third of revenue.

[c] We manipulated the allocation of the compensation pool between the two employees at two levels: a nondiscretionary equal allocation (50 percent to each employee) or a discretionary allocation by the employer.

[d] n = the number of groups in each treatment condition.

[e] The Group(SC × AC) mean square is the appropriate error term for the between-subjects effects. The Period × Group(SC × AC) mean square is the appropriate error term for the within-subjects effects.

[f] We report a one-tailed p-value for the main effect of pool size and pool allocation, given the directional predictions for these effects. All other p-values are two-tailed.

[g] We manipulated Periods as a within-subjects factors with eight levels, 1–8. Participants in all contract conditions completed eight work sessions.

size conditions, while H2b predicted that employee compensation would be greater in the discretionary than the nondiscretionary pool allocation condition. The results in Table 2 support these predictions.

Specifically, Panel A of Table 2 indicates that the mean employee group compensation over all eight periods was 261.23 for the nondiscretionary pool size and 198.60 for the discretionary pool size ($F = 16.07$, $p < 0.01$ as shown in Panel B). Panel A also shows

TABLE 2
Employee Group Compensation (H2a and H2b)

Panel A: Means (Standard Deviations) for Employee Group Compensation over All Periods[a]

Compensation Pool Allocation[c]	Compensation Pool Size[b] Nondiscretionary	Discretionary	Overall
Nondiscretionary	248.44	171.43	208.94
	(50.46)	(81.95)	(78.00)
	n = 19[d]	n = 20	n = 39
Discretionary	273.39	225.77	249.58
	(35.94)	(91.71)	(72.86)
	n = 20	n = 20	n = 40
Overall	261.23	198.60	229.52
	(44.85)	(90.15)	(77.70)
	n = 39	n = 40	n = 79

Panel B: ANOVA on Employee Group Compensation

Source of Variation	SS	df	MS	F[e]	p[f]
Between-Subjects					
Pool Size Condition (SC)	613,244	1	613,244	16.07	< 0.01
Pool Allocation Condition (AC)	248,253	1	248,253	6.51	< 0.01
SC × AC	34,106	1	34,106	0.89	< 0.35
Group(SC × AC)	2,862,249	75	38,163		
Within-Subjects					
Period[g]	182,438	7	26,063	11.20	< 0.01
Period × SC	173,781	7	24,826	10.67	< 0.01
Period × AC	17,064	7	2,438	1.05	< 0.40
Period × SC × AC	15,398	7	2,200	0.95	< 0.48
Period × Group(SC × AC)	1,221,853	525	2,327		

---

[a] Employee groups consist of two employees paired with one employer. Employee group compensation equals the sum of the compensation received by the two employees.

[b] We manipulated the determination of the compensation pool size at two levels: nondiscretionarily set at 2.36 francs (cents) per unit of output or discretionarily set by the employer. In the nondiscretionary case, employees in total received two-thirds of revenue (3.54 per unit of output) and the employer received one-third of revenue.

[c] We manipulated the allocation of the compensation pool between the two employees at two levels: a nondiscretionary equal allocation (50 percent to each employee) or a discretionary allocation by the employer.

[d] n = the number of groups in each treatment condition.

[e] The Group(SC × AC) mean square is the appropriate error term for the between-subjects effects. The Period × Group(SC × AC) mean square is the appropriate error term for the within-subjects effects.

[f] We report a one-tailed p-value for the main effect of pool size and pool allocation, given the directional predictions for these effects. All other p-values are two-tailed.

[g] We manipulated Periods as a within-subjects factors with eight levels, 1-8. Participants in all contract conditions completed eight work sessions.

that the mean group compensation of 249.58 for the discretionary pool allocation condition was greater than that of 208.94 for the nondiscretionary pool allocation condition ($F = 6.51$, $p < 0.01$).

Panel B of Table 2 indicates a statistically significant main effect for period ($F = 11.20$, $p < 0.01$) and a statistically significant period-by-pool size discretion interaction ($F = 10.67$, $p < 0.01$). These interactions are reflected in trends in compensation over the eight periods shown in Panel B of Figure 1. The statistically significant period-by-pool size discretion interaction reflects both decreases in group performance (output) and increases in employers' opportunistic behavior in the discretionary conditions over time. On average, employers in the discretionary pool size conditions paid total employee compensation of 59 percent of revenue, compared to the predetermined 67 percent paid in the nondiscretionary pool size conditions. However, as suggested by the period-by-size discretion interaction, the percentage paid declined over time. For example, in period 1, employers in the two discretionary pool size conditions on average allocated 62 percent of total revenue to employees, while in period 8, they allocated only 51 percent to employees.

Employer's Residual Profit (Hypotheses 3a and 3b)

Hypothesis 3a predicted that the employer's residual profit would be greater in the nondiscretionary size conditions than in the discretionary size conditions. Hypothesis 3b predicted that the employer's residual profit would be greater in the discretionary allocation conditions than in the nondiscretionary allocation conditions. Table 3 and Panel C of Figure 1 present results for the employer's residual profit; these results do not support H3a but do support H3b.

As shown in Panel A of Table 3, employers in the nondiscretionary pool size conditions averaged compensation of 130.65, compared to 139.65 in the discretionary pool size conditions. Contrary to H3a, the main effect of pool size discretion is not statistically significant, as indicated in Panel B ($F = 1.50$, $p < 0.88$). Panel A also indicates that the mean employer compensation was 142.25 in the discretionary pool allocation conditions and 127.98 in the nondiscretionary pool allocation conditions. This difference is statistically significant ($F = 3.68$, $p < 0.03$), consistent with H3b.

As indicated in Panel C of Figure 1, there is considerable variation in employers' compensation across the four contracts and eight periods. This variability is reflected in the statistically significant period-by-pool allocation discretion interaction in Panel B of Table 3 ($F = 2.55$, $p < 0.02$). The results in Panel C of Figure 1 suggest that there is little difference across contracts in employer's residual profit until the last two periods, where profit appears greater in the conditions in which the employer had allocation discretion. We examine this issue further in the supplemental analysis below.

Supplemental Analyses

To provide further insights on our results, we performed two additional analyses. First, we reran all three ANOVAs using only periods 3–6 to eliminate both learning effects in periods 1–2 and end-game effects in periods 7–8. The ANOVAs for group performance and employee compensation showed statistically significant main effects for pool size discretion and pool allocation discretion, similar to results in Tables 1 and 2; however, there were no main or interactive effects related to periods. These analyses indicate that the effects of pool size and allocation discretion documented in Tables 1 and 2 are not simply learning or end-game effects.

The ANOVA for periods 3–6 for employer's residual profit showed no statistically significant main or interactive effects. An ANOVA for periods 1–6 had similar results. Thus, it appears that the statistically significant effect of pool allocation discretion presented in Table 3 arises from results in periods 7–8 and likely represents end-game effects. That is, the employer's ability to differentially allocate compensation among employees appears to benefit the employer primarily near the end of the employment relation.

Second, we performed an ANCOVA for our three dependent variables, adding the group piece rate from the prior period as a covariate. This piece rate was set at 2.36 francs in the two nondiscretionary size conditions, but varied based on the employer's choice in the

TABLE 3
Employer Residual Profit (H3a and H3b)

Panel A: Means (Standard Deviations) for Employer Residual Profit over All Periods[a]

Compensation Pool Allocation[c]	**Compensation Pool Size[b]**		
	Nondiscretionary	**Discretionary**	**Overall**
Nondiscretionary	124.28	131.49	127.98
	(25.27)	(41.87)	(34.53)
	n = 19[d]	n = 20	n = 39
Discretionary	136.70	147.81	142.25
	(17.97)	(41.09)	(31.80)
	n = 20	n = 20	n = 40
Overall	130.65	139.65	135.20
	(22.44)	(41.77)	(33.74)
	n = 39	n = 40	n = 79

Panel B: ANOVA on Employer Residual Profit

Source of Variation	**SS**	**df**	**MS**	**F[e]**	**p[f]**
Between-Subjects					
Pool Size Condition (SC)	13,241	1	13,241	1.50	< 0.88
Pool Allocation Condition (AC)	32,594	1	32,594	3.68	< 0.03
SC × AC	603	1	603	0.07	< 0.80
Group(SC × AC)	664,058	75	8,854		
Within-Subjects					
Period[g]	17,201	7	2,457	1.42	< 0.20
Period × SC	16,467	7	2,352	1.36	< 0.23
Period × AC	30,881	7	4,412	2.55	< 0.02
Period × SC × AC	16,385	7	2,341	1.35	< 0.23
Period × Group(SC × AC)	908,688	525	1,731		

---

[a] Each employer was paired with two employees. An employer's residual profit equals total revenue less employee compensation.

[b] We manipulated the determination of the compensation pool size at two levels: nondiscretionarily set at 2.36 francs (cents) per unit of output or discretionarily set by the employer. In the nondiscretionary case, employees in total received two-thirds of revenue (3.54 per unit of output) and the employer received one-third of revenue.

[c] We manipulated the allocation of the compensation pool between the two employees at two levels: a nondiscretionary equal allocation (50 percent to each employee) or a discretionary allocation by the employer.

[d] n = the number of groups in each treatment condition.

[e] The Group(SC × AC) mean square is the appropriate error term for the between-subjects effects. The Period × Group(SC × AC) mean square is the appropriate error term for the within-subjects effects.

[f] We report a one-tailed p-value for the main effect of pool allocation given the directional predictions for this effect. The p-value for the main effect of pool size equals $(1 - \alpha/2)$, given that the results are in the opposite direction of that predicted in H3a. All other p-values are two-tailed.

[g] We manipulated Periods as a within-subjects factors with eight levels, 1–8. Participants in all contract conditions completed eight work sessions.

discretionary conditions, with an average of 2.09 francs. If employees simply respond to the size of the compensation pool in choosing their work efforts and resulting performance, then it is possible that differences in piece rates between the discretionary and nondiscretionary pool size conditions could lead to our results, rather than employer discretion versus nondiscretion, per se. The ANCOVA tests for this possibility by relating employee performance to the effects of pool size and pool allocation discretion after controlling for any effects of differences in group piece rates.

The prior-period piece rate covariate was statistically significant in all three ANCOVAs. However, the inferences for effects of pool size and allocation discretion did not change between the ANOVAs in Tables 1–3 and the related ANCOVAs. In the ANCOVAs for performance and employee compensation, the main effects for pool size discretion and pool allocation discretion were statistically significant at $p < 0.01$, as in Tables 1 and 2. The ANCOVA for employer residual profit showed a nonstatistically significant effect for size discretion and a statistically significant allocation discretion effect ($p < 0.01$), similar to results in Table 3. Overall, these findings suggest that our results in Tables 1–3 reflect the effects of differences in employer pool size discretion, as opposed to differences in group piece rates between the discretionary and nondiscretionary pool size conditions.

V. SUMMARY AND DISCUSSION

In this paper, we investigate implications of employer discretion over the total amount and allocation of employee compensation for firm output and employer/employee remuneration. We find that a nondiscretionary approach to determining total employee compensation and a discretionary approach to allocating employee compensation increase output and employee pay. The employer's residual profit increases with discretion over pool allocation, but does not increase with a nondiscretionary approach to determining total employee compensation. Additionally, our results indicate that employer and employee opportunism do not completely destroy productive outcomes in the absence of mitigating controls; however, the negative effects of opportunism increase over time, as the employment relationship nears its end.

Our paper has several implications for the role of compensation-based control systems in mitigating potential employer and employee opportunism. First, our paper provides a benchmark for the importance of control systems by documenting the effects of opportunism in the absence of mitigating controls. Our results show that employers and employees do not engage in extreme opportunism in the absence of controls, suggesting that workplace norms lead employees to work and employers to reward this work to their mutual benefit. These findings highlight the need to compare the costs and benefits of compensation systems not only with other control systems, but also with naturally occurring social mechanisms associated with employment settings.[15]

Second, our results support the use of compensation-based control systems since some opportunism exists in their absence. A compensation system based on nondiscretionary compensation pool size and discretionary pool allocation increases firm output. Employees also receive greater pay from such a compensation system; however, that employers do not profit from relinquishing discretion over the size of the compensation pool. This result suggests that employers may decide to retain discretion over aspects of compensation to the possible detriment of total firm output and employee compensation. However, competition among employers for employees could lead to employers relinquishing the right to set the compensation pool size.

---

[15] Some research has found synergies between formal control systems and workplace norms. Specifically, formal control systems can enhance the trust that serves as the foundation for cooperative behavior in business settings (Coletti et al. 2005).

Our paper also complements research on the use of contractible and noncontractible information in compensation contracts. Our empirical results support the analytical conclusions of Baiman and Rajan (1995) regarding the benefits of a contract in which total employee compensation is based on contractible information and the allocation of this compensation is based in part on noncontractible information. Since much of a firm's contractible information emanates from the accounting system, our paper also supports the idea that publicly available accounting information and employer's private information have complementary roles in performance evaluation and compensation.

Future research could provide additional insights on issues examined in this paper. First, as indicated earlier, we made specific choices for compensation parameter values; future research could examine the sensitivity of our results to our compensation parameter values. Specifically, research could examine different piece rates and sharing rules among employers and employees to determine if our results are robust across different parameterizations.

Second, future research could examine other aspects of our design choices, such as anonymous two-employee groups and use of one noisy signal. Employee shirking could increase with larger groups of employees, strengthening the need for compensation-based control systems.[16] In contrast, the effectiveness of discretion over pool allocation might be reduced when employees are not anonymous if employers compensate employees on personal characteristics unrelated to performance (Ittner et al. 2003). Research also could investigate the efficacy of discretionary allocation when employers must arrive at subjective evaluations of employee performance by aggregating several performance signals. Research has documented biases in combining such information (Lipe and Salterio 2000; Prendergast and Topel 1993). Future research could examine whether these biases overcome the benefits of the NDS-DA contract compared to the NDS-NDA contract.

Finally, future research could distinguish between effects of nondiscretion versus commitment in determining the size of the compensation pool. In our experiments, employers who did not have discretion were by definition committed to a compensation contract. However, employers can commit *ex ante* to a contract based on contractible information and still retain discretion over characteristics of the compensation contract, such as its functional form, specific parameters, and length. While we find that employers do not benefit by relinquishing discretion over the size of the compensation pool, it is possible that employers do benefit from committing to a compensation arrangement over which they have discretion to set the terms as suggested by Baiman and Rajan (1995). However, analytic research assumes that employers know employees' utility functions and, therefore, can set contract terms optimally. In practice, the task of understanding *ex ante* precisely what motivates employees may be difficult. Thus, commitment may not benefit an employer compared to the situation of retaining full discretion. Future research on these issues will increase our understanding of how specific characteristics of employment contracts mitigate employee and employer opportunism.

REFERENCES

Axelrod, R. 1984. *The Evolution of Cooperation.* New York, NY: Basic Books.

Baiman, S., and M. V. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70 (October): 557–579.

---

[16] Some public good research, however, has found that individuals' contributions to the common pool actually increase as group size increases (Issac and Walker 1988).

Berg, J., J. Dickhaut, and K. McCabe. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10 (July): 122–142.

Bull, C. 1987. The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics* 102 (February): 147–160.

Chow, C. W. 1983. The effects of job standard tightness and compensation scheme on performance: An exploration of linkages. *The Accounting Review* 58 (October): 667–685.

Coletti, A. L., K. L. Sedatole, and K. L. Towry. 2005. The effect of control on trust and cooperation in collaborative environments. *The Accounting Review* 80 (April): 477–500.

Demski, J. 1997. *Managerial Uses of Accounting Information.* Boston, MA: Kluwer Academic Publishers.

Fehr, E., S. Gächter, and G. Kirchsteiger. 1997. Reciprocity as a contract enforcement device. *Econometrica* 65 (July): 833–860.

———, and ———. 2000. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14 (Summer): 159–181

Fisher, J. G., S. A. Peffer, and G. B. Sprinkle. 2003. Budget-based contracts, budget levels, and group productivity. *Journal of Management Accounting Research* 15: 51–74.

Isaac, R. M., and J. M. Walker. 1988. The assurance problem in a laboratory market. *Quarterly Journal of Economics* 103 (July): 179–200.

Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (July): 725–758.

Ledyard, J. O. 1995. Public goods: A survey of experimental research. In *Handbook of Experimental Economics,* edited by H. H. Kagel, and A. E. Roth, 111–194. Princeton, NJ: Princeton University Press.

Lipe, M. G., and S. E. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75 (July): 283–298.

Milgrom, P. R., and J. Roberts. 1992. *Economics, Organization & Management.* Upper Saddle River, NJ: Prentice Hall.

Prendergast, C., and R. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37 (April): 355–365.

———. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37 (March): 7–63.

Sprinkle, G. B. 2003. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (February–April): 287–318.

Zimmerman, J. L. 2003. *Accounting For Decision Making and Control.* Boston, MA: McGraw-Hill.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 585–612

Auditor Tenure and Perceptions of Audit Quality

Aloke Ghosh
U.S. Securities and Exchange Commission
and
Baruch College—The City University of New York

Doocheol Moon
State University of New York at Old Westbury

ABSTRACT: We analyze how investors and information intermediaries perceive auditor tenure. Using earnings response coefficients from returns-earnings regressions as a proxy for investor perceptions of earnings quality, we document a positive association between investor perceptions of earnings quality and tenure. Further, we find that the influence of reported earnings on stock rankings becomes larger with extended tenure, although the association between debt ratings and reported earnings does not vary with tenure. Finally, we find that the influence of past earnings on one-year-ahead earnings forecasts becomes greater as tenure increases. In general, our results are consistent with the hypothesis that investors and information intermediaries perceive auditor tenure as improving audit quality. One implication of our study is that imposing mandatory limits on the duration of the auditor-client relationship might impose unintended costs on capital market participants.

Keywords: *auditor tenure; auditor independence; audit quality; earnings quality; mandatory auditor rotation; capital market perceptions.*

Data Availability: *All data used in this study are available from public sources.*

We thank Ashiq Ali, Mark Beasley, Marti Benis, Donal Byard, Masako Darrough, Paquita Davis-Friday, Anita Dennis, Zeng Deng, Rajib Doogar, Amy Edwards, Mary Greenawalt, Zhaoyang Gu, Larry Harris, Prem Jain, Joe Kerstein, Ying Li, Steve Lilien, Steve Lustgarten, Darius Miller, two anonymous referees, Srini Sankaraguruswamy, Bharat Sarath, Dave Smith, Jonathan Sokobin, Jan Sweeney, Tony Tinker, Joe Weintrop, and Jimmy Ye for their helpful comments. This paper was previously titled: "Does Auditor Tenure Impair Audit Quality?" We acknowledge financial support from the Robert Zicklin Center for Corporate Integrity of Baruch College.

The Securities and Exchange Commission disclaims responsibility for any private publication or statement by any SEC employee or commissioner. This paper expresses the authors' views and does not necessarily reflect those of the Commission, the commissioners, or other members of the staff.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted August 2002
Accepted August 2004

I. INTRODUCTION

The recent rise in accounting irregularities has reopened questions about auditor tenure, independence, and audit quality (Bricker 2002).[1] Recent studies provide valuable insights into the debate surrounding auditor tenure by examining the association between tenure and (1) accounting accruals, (2) analysts' forecast errors, and (3) the cost of debt. Myers et al. (2003) conclude that longer auditor tenure constrains managerial discretion with accounting accruals, which suggests high audit quality. Johnson et al. (2002) also find that accruals are larger and less persistent for firms with short auditor tenure relative to those with medium or long tenure. Using credit spreads between bond yields and matched Treasury yields as the cost of debt, Mansi et al. (2004) find that the cost of debt declines with longer tenure, which suggests bondholders perceive audit quality as improving with extended tenure. In contrast, Davis et al. (2002) conclude that audit quality declines with extended tenure because, as tenure increases, client firms have greater reporting flexibility and earnings forecast errors decline.

This study focuses on how investors and information intermediaries perceive auditor tenure. Using earnings response coefficients from returns-earnings regressions as a proxy for investor perceptions of earnings quality, we analyze whether investors perceive earnings quality as being affected by tenure. Given the importance of information intermediaries who receive and process financial information for investors, we also analyze whether independent rating agencies and financial analysts incorporate the potential effects of tenure on earnings quality. Specifically, we examine whether tenure affects the relationship between reported earnings and (1) stock rankings, (2) debt ratings, and (3) analysts' earnings forecasts.

A key distinction between our study and those of Davis et al. (2002) and Mansi et al. (2004) is the difference in the research design. We examine whether the extent to which analysts rely on past reported earnings to predict future earnings varies with tenure, whereas Davis et al. (2002) explore the association between forecast errors and tenure. Our research design might be better suited because it is difficult to draw unambiguous inferences about analysts' perceptions from an association between forecast errors and tenure. Lower forecast errors with longer tenure might suggest that earnings quality is perceived as improving with tenure because earnings are more predictable. However, it could also suggest lower earnings quality if managers increasingly guide earnings forecasts as auditor tenure lengthens. Further, Mansi et al. (2004) investigate the influence of tenure on the cost of debt, whereas we examine whether the influence of reported earnings on debt ratings varies with tenure. Although Mansi et al. (2004) examine the effect of tenure on debt ratings, their primary aim is to purge (orthogonalize) the information effects of tenure and other control variables related to debt ratings and ultimately examine the insurance role of tenure on the cost of debt. In contrast, our emphasis is on the informational role of tenure. Our fundamental objective is to provide insights into any changes in the perceived credibility of reported earnings with extended auditor tenure.

Emphasis on capital market perceptions of independence and audit quality is consistent with the Financial Accounting Standards Board's (FASB) conceptual framework for financial reporting and principles of auditor independence (Carmichael 1999). According to the

---

[1] Regulators express concerns that pressure to retain client firms and the "comfort level" created between auditors and management over time impair auditor independence, which adversely affects audit quality (U.S. Government Accounting Office [GAO] 2003). The accounting profession, on the other hand, claims that the likelihood of audit failures is greater during the initial period of an auditor-client relationship because of lack of information about client-specific risks (PricewaterhouseCoopers 2002).

Statement of Financial Accounting Concepts (SFAC No. 1, FASB 1978), "financial statements are often audited by independent accountants for the purpose of enhancing confidence in their reliability." The AICPA (1994) also acknowledges the importance of considering investor perceptions of auditor independence. A former chairman of AICPA, Elliott (2000) says "[The AICPA] believe[s] that appearances are very important and capital markets require confidence in financial statements and audit reports, and the member firms of the AICPA are basing their business of auditing on their reputations, and that is heavily affected by appearance."

Other things remaining constant, audited financial statements are less reliable (or of lower perceived quality) for investment and credit decisions if users of financial statements view lengthy tenure as having an adverse effect on auditor independence and audit quality. Alternatively, investors and information intermediaries are more likely to rely on reported accounting numbers if they perceive that greater auditor expertise from longer tenure improves independence and audit quality. We assert that the relationship between perceptions of earnings quality and auditor tenure provides insights into how capital market participants view auditor tenure as affecting audit quality.

We use earnings response coefficients (ERCs) from contemporaneous returns-earnings regressions to measure investor perceptions of earnings quality. After controlling for the other determinants of ERC—such as the age of the firm, the quality of auditors, growth, earnings persistence, earnings volatility, systematic risk, firm size, financial leverage, and regulatory environment—we find that the magnitude of the ERC increases as the auditor-client relationship lengthens. One concern is that if the market anticipates a portion of current earnings more than one-year-ahead of the earnings release, then the estimated ERC might be biased downward (Kothari 1992; Kothari and Sloan 1992). Additionally, if the market is more likely to anticipate current earnings for firms with longer tenure, then the bias might be associated with tenure. Consistent with the premise that prices lead earnings, the estimated ERC for firms with extended tenure is larger when we increase the returns measurement window or when we use non-market metrics, such as total assets, to deflate earnings. Thus, the ERC-based findings are consistent with the hypothesis that investors perceive auditor tenure as enhancing earnings quality.

Further, controlling for the other determinants of stock rankings, we find that the influence of reported earnings on Standard & Poor's (S&P) common stock rankings becomes larger with extended tenure. In contrast, the association between S&P debt ratings and reported earnings does not vary with tenure. Thus, our results provide modest evidence that independent rating agencies perceive reported earnings as being more reliable for firms with longer tenure.

Finally, after controlling for factors that affect analysts' earnings forecasts, we find that the influence of past earnings on one-year-ahead earnings forecasts becomes larger over extended auditor-client relationships. All else equal, analysts are more likely to rely on reported earnings to predict future earnings with longer tenure. Thus, the results from analyst earnings forecasts also suggest that analysts perceive earnings quality as improving with longer auditor tenure.

While our results are based on a large sample of firms spanning 11 years, there is one concern with our dataset. If auditor tenure is endogenous to audit quality, then the results are also consistent with an alternative hypothesis that auditor turnover is high for firms with low earnings quality. In other words, high-quality auditors might terminate engagements with client firms that prefer low-quality financial statements (DeFond and Subramanyam 1998). We address potential concerns about frequent auditor turnover by constructing a subsample for which the auditor-client relationship lasts for at least five years, as in Myers et al. (2003). When we use this subsample, our inferences remain unchanged for all our tests.

In sum, our results are generally consistent with the hypothesis that reported earnings are perceived as being more reliable as auditor tenure increases. One implication of our results is that capital market participants view longer auditor tenure as having a favorable impact on audit quality. Our results suggest that imposing mandatory limits on the duration of the auditor-client relationship might impose unintended costs on capital markets. However, results from a regime without auditor term limits may not be applicable to a regulated environment because of differences in economic incentives for both auditors and clients.

The rest of the paper is organized as follows. Section II establishes the links between tenure, independence, and perceptions of audit quality. Section III discusses the research design, and Section IV describes the sample selection procedure. Section V reports the results of the association between tenure and perceptions of earnings quality. Section VI concludes the paper.

II. TENURE, INDEPENDENCE, AND PERCEPTIONS OF AUDIT QUALITY

Independent auditors are considered the "gatekeepers" of the public securities markets (SEC 2001, III.A). However, the recent rash of accounting irregularities has led many to question auditor independence (*Wall Street Journal* 2002a, 2002b). One perception is that auditors are more likely to agree with managers on important reporting decisions as the length of the audit engagement increases (Ryan et al. 2001; Farmer et al. 1987). Therefore, imposing mandatory limits on auditor tenure is expected to improve audit quality by reducing client firms' influence over auditors (Turner 2002; Brody and Moscove 1998; SEC 1994; AICPA 1978; U.S. Senate 1977; Mautz and Sharaf 1961).

An opposing viewpoint is that problem audits occur more frequently for newer clients because auditors have less information about these firms (AICPA 1992). Client-specific knowledge of items such as operations, accounting system, and internal control structure is crucial for auditors to detect material errors and misstatements. In particular, Johnson et al. (2002) argue that lack of adequate client-specific knowledge during the early years of engagement decreases the likelihood of detecting material errors and misstatements. As the auditor-client relationship lengthens, firm-specific expertise allows auditors to rely less on managerial estimates and become more independent of management (Solomon et al. 1999).

The crux of the debate rests on how tenure affects auditor independence. Proponents of mandatory auditor rotation claim that lengthy auditor tenure erodes independence, which in turn impairs audit quality. Others argue that independence and audit quality increase with longer tenure because of improved auditor expertise from superior client-specific knowledge.[2] Since independence is not observable, regulators, practitioners, and academics often rely on the appearance dimension to define auditor independence (Dopuch et al. 2003; Kinney 1999).

Our study investigates whether capital market participants perceive longer tenure as affecting audit quality. Insights from a market-based approach are important along at least two dimensions. First, academics often emphasize the need to understand capital market perceptions of auditor independence and audit quality because ultimately the value of auditing services depends on perceptions of independence (Dopuch et al. 2003; Shockley

---

[2] Expert knowledge gained through years of on-the-job experience increases the likelihood that auditors will detect errors in financial statements (Ashton 1991; Libby and Frederick 1990). In contrast, industry specialization, another frequently used proxy for auditor expertise, is based on training and practical experience gained from auditing in a particular industry (Gramling and Stone 2001; Solomon et al. 1999; Hogan and Jeter 1999; Craswell et al. 1995).

1981). Second, regulatory institutions such as the FASB (SFAC No. 1) and SEC (2000) emphasize the importance of capital market perceptions of auditor independence. A perception that the auditors' work is more objective and independent inspires greater confidence in auditor opinion, which increases the perceived reliability or quality of reported accounting numbers (Ryan et al. 2001; Elliott and Jacobson 1998). Independent auditors increase the reliability of financial statements because (1) they are more likely to prevent or detect and correct material misstatements/omissions, and (2) they ensure that financial statements comply with generally accepted accounting principles (Carmichael 1999). Therefore, to the extent that capital market participants view tenure as improving independence and audit quality, financial statements are perceived as more reliable for financial decisions as tenure lengthens.

In this study, we focus on the perceptions of investors because they are the principal users of financial statements. In its Conceptual Framework, the FASB defines "quality" in relation to the usefulness of financial statements to investors and links "usefulness" in turn to constructs such as relevance and reliability (SFACs No. 1 [FASB 1978] and No. 2 [FASB 1980]). To draw inferences about investors' perceptions of earnings quality, researchers tend to use stock-market-based metrics such as earnings response coefficients from regressions of returns on earnings (Schipper and Vincent 2003; Warfield et al. 1995). Prior studies document that investors pay a larger premium for "high-quality" earnings because high-quality earnings are viewed as sustainable (Schipper and Vincent 2003; Teoh and Wong 1993). Thus, examining the influence of auditor tenure on the pricing of earnings is likely to provide valuable insights into investors' views of the association between the length of the auditor-client relationship and earnings quality.

Given that information intermediaries constitute an integral part of the capital market by providing stock recommendations, debt ratings, and earnings forecasts (Lang and Lundholm 1996), we also analyze how independent rating agencies and financial analysts view auditor tenure. Hunt (2002) states that independent rating agencies provide information about the creditworthiness of issuers and that credit ratings play a significant role in investment decisions. Extant research finds links between earnings and debt ratings/stock rankings issued by independent rating agencies (Bhojraj and Sengupta 2003; Ziebart and Reiter 1992; Van Horne 1992; Kaplan and Urwitz 1979), suggesting that perceptions of earnings quality could be an important input in determining rankings/ratings. Similarly, financial analysts also play a prominent role as information intermediaries in capital markets because of their ability to incorporate value-relevant information in their published reports, which impacts security prices (Francis and Soffer 1997; Schipper 1991; Lys and Sohn 1990; Brown et al. 1987). Prior research finds that analysts rely on earnings releases to estimate future earnings (Kasznik and McNichols 2002; Barron et al. 2002; Stickel 1989), which suggests that the extent to which analysts depend on reported earnings to make earnings forecasts might vary with perceptions of earnings quality.

Extending this line of research, we examine the association between auditor tenure and (1) stock rankings, (2) debt ratings, and (3) analysts' forecasts of earnings per share. Additionally, we analyze how auditor tenure affects the association between reported earnings and rankings, ratings, and earnings forecasts. If auditor tenure is perceived as enhancing earnings quality, then, all else equal, the influence of reported earnings on rankings/ratings and earnings forecasts is expected to become larger with longer auditor tenure because reported earnings are viewed as more informative about future earnings. The converse is true if information intermediaries perceive longer auditor tenure as eroding earnings quality.

Based on the association between our proxy measures for perceptions of earnings quality and auditor tenure, we infer how investors and information intermediaries view auditor tenure as affecting audit quality. A positive relationship between the proxies for perceptions of earnings quality and auditor tenure is consistent with the hypothesis that longer tenure is perceived as improving independence and audit quality. In contrast, a negative relationship is consistent with the hypothesis that lengthy tenure is viewed as eroding independence and audit quality.

III. RESEARCH DESIGN

We use the following basic regression framework to analyze whether investors, independent rating agencies, and financial analysts perceive earnings quality as being affected by auditor tenure:

$$\begin{aligned} \textit{Dependent variable} = {} & \alpha + \beta_1 E + \beta_2 \Delta E + \beta_3 E^{*}\textit{Tenure} + \beta_4 \Delta E^{*}\textit{Tenure} \\ & + \beta_5 \textit{Tenure} + \sum_{j=1}^{9} \beta_{6+2(j-1)} E^{*}\textit{Control variable}_j \\ & + \sum_{j=1}^{9} \beta_{7+2(j-1)}\, \Delta E^{*}\textit{Control variable}_j \\ & + \sum_{j=1}^{9} \beta_{23+j} \textit{Control variable}_j + \varepsilon. \end{aligned} \tag{1}$$

E and ΔE are reported earnings and changes in reported earnings, respectively. *Tenure* is the duration of the auditor-client relationship (calculated using Compustat #149) in years, starting from 1982 (audit firm mergers are considered as a continuation of the prior auditor). Subscript *j* represents the number of control variables (1 to 9). Each of the control variables is interacted with $E/\Delta E$ and is also included as a separate independent variable. Details on the dependent and control variables are provided in subsequent subsections.

The sum of earnings levels and changes coefficients ($\beta_1 + \beta_2$) or the "earnings response coefficient" (ERC) is our proxy for capital markets' perceptions of earnings quality. Our interest is in the sum of the *E*Tenure* and *ΔE*Tenure* coefficients ($\beta_3 + \beta_4$). If investors, rating agencies, and analysts perceive earnings quality as improving (declining) with longer auditor tenure, $\beta_3 + \beta_4$ is expected to differ from zero.

Perceptions of Investors and Auditor Tenure

We measure investor perceptions of earnings quality using 12-month (ending three months after the fiscal year-end) cumulative market-adjusted returns (*CAR*) as the dependent variable in Equation (1). Market-adjusted returns are the difference between raw returns and value-weighted CRSP market returns.

We include earnings changes and earnings levels in the same regression because including both increases the explanatory power and magnitude of earnings response coefficients when earnings contain both transitory and permanent components (Easton and Harris 1991; Ali and Zarowin 1992). *E* is income before extraordinary items (Compustat #18) and ΔE is the difference between income before extraordinary items for the current year and that of last year. Both *E* and ΔE are deflated by market value of equity (Compustat #25 × #199) at the beginning of the year.

We include various control variables because the ERC is associated with other firm characteristics (Warfield et al. 1995; Dhaliwal and Reynolds 1994; Collins and Kothari 1989), which, in turn, might be associated with tenure. The control variables are defined as follows: *FirmAge*, computed using the beginning and end dates as reported in CRSP, measures the number of years that the firm has been publicly traded as of the fiscal year-end; *Big4* is an indicator variable that equals 1 when the client's auditor is a large accounting firm (Compustat #149); *Growth* is the sum of the market value of equity (Compustat #25 × #199) and the book value of debt (Compustat #9 + #34) scaled by the book value of total assets (Compustat #6); *Persistence* (*Volatility*) is the first-order autocorrelation (standard deviation) of income before extraordinary items per share (Compustat Quarterly #8/ Quarterly #61) for the past 16 quarters;[3] *Beta* is systematic risk computed using the past 60 monthly stock returns; *Size* is the logarithmic transformation of the fiscal year-end market value of equity (Compustat #25 × #199) of the prior year; *Leverage* is the ratio of total debt (Compustat #9 + #34) to total assets (Compustat #6); and *Regulation* is an indicator variable that equals 1 for firms in a regulated industry with two-digit standard industry classification codes between 40 and 49 or between 60 and 63.

We include *FirmAge* for two reasons: (1) older firms are more likely to be stable with less information asymmetry problems, which suggests higher ERCs; and (2) *Tenure* and *FirmAge* are positively correlated. We control for *Big4* because large auditors are generally associated with high-quality audits (Becker et al. 1998; Teoh and Wong 1993). The inclusion of *Growth, Persistence, Volatility*, and *Beta* is primarily motivated by valuation considerations (Warfield et al. 1995). The inclusion of *Size* is motivated by the political cost theory: managers of large, politically sensitive firms are more likely to exploit the latitude in accounting to reduce political costs, which affects earnings quality. We include *Leverage* because of contracting considerations: firms with high leverage are more likely to use the latitude in accounting to avoid possible debt-covenant violations (DeFond and Jiambalvo 1994). Finally, the quality of earnings is affected in a regulated environment (*Regulation*) because managers have limited scope for opportunistic behavior (Warfield et al. 1995).

Perceptions of Information Intermediaries and Auditor Tenure

Independent Rating Agencies

Our analysis of how tenure affects independent rating agencies' perceptions of earnings quality is based on estimates from Equation (1) using Standard & Poor's (S&P) common stock rankings (*Stock Rankings*) and senior debt ratings (*Debt Ratings*) as dependent variables. *Stock Rankings* represent numerical values of 1 to 7, corresponding to S&P common stock rankings (Compustat #282). A value of 1 is assigned if a firm's common stock ranking is rated as A+ and as S&P common stock rankings decline, the numerical value increases by 1.[4] Similarly, *Debt Ratings* are assigned a value of 1 if a firm's S&P senior debt (Compustat #280) is rated as AAA. As S&P debt ratings decline from AAA (the highest) to D (payment default), the numerical value increases by 1.

We include *Growth, Volatility, Beta, Size*, and *Leverage* to control for cross-sectional differences in firm quality and riskiness (Bhojraj and Sengupta 2003; Van Horne 1992; Ziebart and Reiter 1992; Kaplan and Urwitz 1979). Growing firms with high earnings volatility tend to be more risky, *Beta* controls for both operating and financial risk, large firms tend to be less risky, and firms with higher leverage have greater financial risk. We also include *FirmAge, Big4, Persistence*, and *Regulation* because the association between reported earnings and rankings/ratings might depend on the firm's life-cycle, whether firms use big audit firms, whether earnings are persistent, and whether firms are in a regulated environment.

---

[3] Since *Persistence* and *Volatility* measures are computed using time-series income per share numbers, we adjust for stock-splits and stock dividends that occur subsequent to the end of a given period using the adjustment factor reported in Compustat (Quarterly #17).

[4] For instance, values 2 through 7 correspond to the S&P common stock rankings of A, A–, B+, B, B–, and C, respectively. We do not include S&P common stock rankings of D (in reorganization) and LIQ (liquidation) for our sample because of going-concern issues.

In a recent study, Mansi et al. (2004) find that longer tenure lowers the rate of return required by bondholders and that the impact is larger for "information-sensitive" securities such as non-investment grade bonds (debt ratings less than BBB−). Therefore, we also separately estimate the effects of tenure on investment and non-investment grade debt ratings.

Financial Analysts

Finally, our analysis of whether tenure influences financial analysts' perceptions of earnings quality is based on the estimates from Equation (1) using earnings forecasts ($FEPS_t$) as the dependent variable. $FEPS_t$ is the consensus analyst forecasts of annual earnings per share from the Institutional Brokers Estimation System (I/B/E/S) database. Our proxy for consensus forecasts is the mean one-year-ahead forecast for year t issued immediately following the earnings announcement for year $t-1$ (earnings announcement dates are obtained from Compustat).[5] We restrict the forecasts to those issued immediately following earnings announcements, as in Kasznik and McNichols (2002) and Barron et al. (2002), because (1) any change in analysts' perceptions of earnings quality is more likely to be updated following earnings releases (Barron et al. 2002), (2) it excludes stale forecasts since forecast revisions are more common following earnings announcements (Stickel 1989), and (3) it conditions the forecasts on the same set of publicly disclosed information.

Reported earnings per share (EPS_{t-1} and ΔEPS_{t-1}) are used as measures for E and ΔE in Equation (1). EPS_{t-1} is annual earnings per share reported for year $t-1$, and ΔEPS_{t-1} is the absolute change in earnings per share for year $t-1$ defined as the difference in annual earnings per share in year $t-1$ and that in year $t-2$ ($|EPS_{t-1} - EPS_{t-2}|$), both obtained from the I/B/E/S database. As in Barron et al. (2002), we include the absolute value of changes in earnings because a larger earnings surprise tends to be temporary, which reduces the potential usefulness of past earnings in predicting future earnings.

Since prior studies find that analysts' incentives to acquire information about future earnings are affected by firm characteristics (DeFond and Hung 2003; Barron et al. 2002; Lang and Lundholm 1996), we include the following control variables: *FirmAge* because older firms are more likely to be stable and consequently earnings might be easier to predict; auditor type (*Big4*) because analysts might perceive earnings quality as improving for firms with big auditors; *Growth* because high-growth firms tend to generate greater demand for private information, thereby reducing the reliance on reported earnings (Barron et al. 2002; Lang and Lundholm 1996); *Volatility* because analysts are more likely to attach lower importance on reported earnings for firms with higher earnings volatility (DeFond and Hung 2003); risk using a market-based measure (*Beta*) and a balance-sheet-based measure (*Leverage*); *Size* because prior studies find that firm size is associated with risk and the information environment, which affects earnings predictability (DeFond and Hung 2003; Barron et al. 2002); *Regulation* because the predictability of earnings might vary between industries (O'Brien 1990); and the number of analysts (*Analysts*) providing annual earnings per share forecasts for a firm (Lang and Lundholm 1996). All the control variables including *Tenure* are measured at the end of year $t-1$.

---

[5] We prefer consensus forecasts to individual forecasts because prior research finds that consensus forecasts are more accurate than the individual forecasts underlying the consensus. One explanation for this result is that the forecast errors made by individual analysts are less than perfectly correlated and therefore in the process of aggregation, the individual forecast errors offset each other (see Brown et al. [1985] for related papers). The results are similar when we use either the mean or the median forecast.

In a recent study, Davis et al. (2002) find that forecast errors decline as tenure increases.[6] However, it is difficult to draw unambiguous inferences about perceptions of analysts from an association between forecast errors and tenure. Lower forecast errors might suggest that analysts perceive earnings quality as improving with tenure because earnings are more predictable. Alternatively, lower forecast errors with longer tenure might suggest lower earnings quality if managers are more likely to guide earnings forecasts as auditor tenure lengthens.[7] Hence, we do not focus on earnings forecast attributes such as forecast errors and dispersion.

IV. DATA

We construct a sample from the list of publicly traded firms in the 2001 Compustat annual files (active and research). Since the 2001 Compustat files cover 20 years of data, and financial data are available from 1982, auditor tenure is 1 for the first year by construction. The analysis begins with 1990 to provide some variation for auditor tenure and excludes the year 2001 because of possible missing and incomplete data.[8] Stock return and firm age data are obtained from the 2001 CRSP files. We get earnings forecast data from the 2001 Institutional Brokers Estimation System (I/B/E/S) Summary Estimates.

We impose the following restrictions on the sample: (1) we delete the top and bottom 1 percent of observations for the level of earnings (E), changes in earnings (ΔE), annual earnings per share (EPS), and the absolute change in annual earnings per share (ΔEPS); (2) we remove all observations with the absolute value of cumulative market-adjusted returns (CAR) greater than 100 percent; and (3) we also winsorize the top and bottom 1 percent of observations for *Growth*, *Persistence*, *Volatility*, *Beta*, and *Leverage*.[9] This sample selection procedure results in a maximum of 38,794 observations for the "full" sample over the years 1990 through 2000.

Some concerns remain with the full sample. Firms might frequently switch auditors either because of "opinion shopping" (SEC 1988) or because of auditors' preference for conservative accounting choices (DeFond and Subramanyam 1998). Alternatively, high-quality auditors might end engagements with clients that prefer low quality of financial reporting. We avoid some of these biases by constructing a subsample in which the auditor-client relationship lasts for at least five years. For auditor-client relationships lasting less than five years, we delete all of the firm-year observations. The maximum number of observations for the "restricted" sample is 35,826 firm-years.[10]

---

[6] Researchers frequently use forecast errors and dispersion among forecasts as proxies for the uncertainty among analysts about the quality of accounting information (Hope 2003; Duru and Reeb 2002; Ashbaugh and Pincus 2001; Barron et al. 1998; Lang and Lundholm 1996).

[7] Managers have strong incentives to avoid negative earnings surprises because of the large stock-price reactions to negative earnings (see Ghosh et al. 2005; Bartov et al. 2002; Myers and Skinner 2002; Skinner and Sloan 2002; Barth et al. 1999). Firms can avoid earnings surprises by manipulating earnings to meet expectations or by guiding analyst expectations to avoid overly optimistic forecasts. Although existing studies suggest that earnings management declines with longer auditor tenure (Myers et al. 2003; Johnson et al. 2002), we do not know whether tenure affects managers' propensity to guide analysts' expectations.

[8] Our computation of *Tenure* will understate the actual length of auditor tenure because *Tenure* is 1 for firms in 1982 even if the auditor was the same for the prior years. However, when we replicate our analysis computing *Tenure* from 1974 using information about the client's auditor as reported in Compustat backfiles, similar to Mansi et al. (2004), our conclusions remain unaffected.

[9] The use of alternative cut-off points such as 0.5 percent also yields similar results, but it retains extreme observations. For instance, the minimum of E is –541 percent and the maximum of ΔE is 629 percent when we delete the top and bottom 0.5 percent. We delete observations with absolute *CAR*s greater than 100 percent to reduce the impact of major corporate restructuring activities such as mergers, bankruptcies, spin-offs, etc. However, when we delete the top and bottom 1 percent of *CAR*, we get similar results. The results on tenure are unaffected when we truncate the top and bottom 1 percent of observations for the control variables.

Table 1 reports summary statistics for both the full sample (Panel A) and the restricted sample (Panel B). We report the mean, minimum, first quartile, median, third quartile, and maximum for *Tenure*, *FirmAge*, *CAR*, *E*, ΔE, *Stock Rankings*, *Debt Ratings*, and *FEPS*. Auditor tenure is longer for the restricted sample than for the full sample; the mean/median

TABLE 1
Descriptive Statistics

Variables	Mean	Min	25%	50%	75%	Max	n
Panel A: Full Sample							
Tenure (years)	8.549	1.000	5.000	8.000	12.000	19.000	38,794
FirmAge (years)	17.597	3.417	7.000	12.167	23.833	75.000	38,794
CAR	−0.060	−0.991	−0.343	−0.108	0.158	0.999	38,794
E	−0.007	−2.795	−0.022	0.048	0.083	0.523	38,794
ΔE	0.006	−1.418	−0.032	0.006	0.034	2.815	38,794
Stock Rankings	5.096	1.000	4.000	5.000	6.000	7.000	25,745
Debt Ratings	9.147	1.000	6.000	9.000	12.000	23.000	9,462
FEPS ($)	1.045	−4.140	0.450	0.890	1.490	7.250	16,417
Panel B: Restricted Sample							
Tenure (years)	9.096	1.000	6.000	9.000	12.000	19.000	35,826
FirmAge (years)	18.004	3.417	7.000	12.750	24.000	75.000	35,826
CAR	−0.054	−0.989	−0.332	−0.102	0.161	0.999	35,826
E	0.000	−2.795	−0.013	0.050	0.084	0.523	35,826
ΔE	0.006	−1.418	−0.030	0.006	0.033	2.804	35,826
Stock Rankings	5.055	1.000	4.000	5.000	6.000	7.000	24,216
Debt Ratings	9.087	1.000	6.000	9.000	12.000	23.000	9,060
FEPS ($)	1.049	−4.140	0.450	0.890	1.500	7.250	15,699

The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. *FirmAge* measures the number of years that the firm has been publicly traded as of the fiscal year-end. *CAR* is cumulative market-adjusted returns for the 12-month period ending three months after the fiscal year-end. Market-adjusted returns are the difference between raw returns and value-weighted CRSP market returns. *E* is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. *Stock Rankings* and *Debt Ratings* represent numerical values of S&P common stock rankings and S&P senior debt ratings, respectively. *Stock Rankings* (*Debt Ratings*) are assigned a value of 1 if a firm's S&P common stock (senior debt) is rated as A+ (AAA), and as the S&P common stock rankings (debt ratings) decline, the numerical value increases by 1. *FEPS* is the mean annual one-year-ahead earnings per share forecasts for year t issued immediately following the earnings announcement for year $t-1$.

---

[10] Although we use the Myers et al. (2003) methodology to construct the restricted sample, the percentage of observations deleted from the full sample is much smaller in our study. One primary reason for the difference is that we include many more variables in our analyses, which reduces the size of our full sample. Thus, fewer observations are lost from requiring that the auditor-client relationship lasts for at least five years in our study relative to those in Myers et al. (2003).

Tenure for the full sample is 8.5/8 years, while the corresponding number for the restricted sample is 9.1/9 years. The mean *FirmAge* for the full sample is 17.6 years, compared to a mean of 18.0 years for the restricted sample. *FirmAge* ranges between 3.4 years and 75 years for the full and restricted samples.

The summary statistics for *CAR*, *E*, and ΔE indicate few differences across the two samples. The mean/median *CAR* for the full sample is −0.060/−0.108, while that for the restricted sample is −0.054/−0.102.[11] The mean/median *E* for the full sample is −0.007/ 0.048, compared to a mean/median of 0.000/0.050 for the restricted sample. The mean/ median ΔE is identical for the full and restricted samples (0.006).

Similarly, there are almost no differences in *Stock Rankings*, *Debt Ratings*, and earnings forecasts (*FEPS*) between the two samples. The average firm in both samples has a B S&P common stock ranking (a numerical score of 5), and a BBB S&P senior debt rating (a numerical score of about 9). The mean/median one-year-ahead analysts' earnings per share forecasts (*FEPS*) is $1.05/$0.89 and $1.05/$0.89 for the full and restricted samples, respectively.

Table 2 reports the Pearson correlation (ρ) matrix between *CAR*, *E*, ΔE, *Stock Rankings*, *Debt Ratings*, *FEPS*, *Tenure*, and *FirmAge* for the full sample. The high correlation between *Stock Rankings* and *Debt Ratings* ($\rho = 0.686$) suggests that stock rankings and debt ratings issued by independent rating agencies contain significant overlapping information. However, low correlations across variables that measure investor, rating agency, and financial analyst perceptions indicate that analysis of each market participant's perceptions provides distinct insights into the overall perceptions of the quality of accounting information. Finally, the correlation between *Tenure* and *FirmAge* is positive and statistically significant ($\rho = 0.337$, p-value = 0.001), which underscores the need to control for *FirmAge* when analyzing *Tenure*.

V. RESULTS

All the regression results reported in Tables 3 to 7 are based on yearly regressions from 1990 to 2000. The reported coefficients are the average of 11 yearly coefficients, and the corresponding t-statistics are computed by comparing the average coefficient to its time-series standard error. We prefer this approach to a pooled time-series cross-sectional estimation because the standard errors in the latter approach are not adjusted for the correlation of regression residuals across firms, which lead to inflated t-statistics. The estimation procedure of year-by-year cross-sectional regressions and the use of average slopes and their time-series standard errors to draw inferences allow for residual cross-correlation (for details, see Fama and French 2000).

Perceptions of Investors and Auditor Tenure

In this subsection, we report the findings of how auditor tenure affects the returns-earnings association. Our interest is in the sign and magnitude of the sum of the coefficients on E**Tenure* and ΔE**Tenure* ($\beta_3 + \beta_4$). Consistent with prior research, we find that reported earnings (*E* and ΔE) are significantly positively associated with returns (*CAR*) in Table 3. The sum of the coefficients on *E* and ΔE, or ERC, is 0.507 (t-statistic = 10.30) in the first

---

[11] Prior studies also find negative mean *CAR*: (1) DeFond and Hung (2003) report that the mean *CAR* measured *over fifteen months* for a sample of 21,908 observations from 1993 to 1999 is −0.05 (−0.04) for firms with (without) I/B/E/S cash flow forecasts; (2) Subramanyam and Wild (1996) find that the mean *CAR over one quarter* for a sample of 25,160 observations from 1981 to 1990 is −0.01; and (3) Ali and Zarowin (1992) report that the mean *12-month CAR* from 1970 to 1985 is −0.05 (−0.04) for the transitory (permanent) group with 4,355 (7,219) observations.

TABLE 2
Pearson Correlation Matrix

Variables	1	2	3	4	5	6	7	8
1. *CAR*	1.000							
2. *E*	0.262	1.000						
	(0.001)							
3. ΔE	0.194	0.368	1.000					
	(0.001)	(0.001)						
4. *Stock Rankings*	−0.118	−0.266	0.068	1.000				
	(0.001)	(0.001)	(0.001)					
5. *Debt Ratings*	−0.185	−0.341	−0.028	0.686	1.000			
	(0.001)	(0.001)	(0.004)	(0.001)				
6. *FEPS*	−0.002	0.021	−0.003	−0.007	−0.080	1.000		
	(0.802)	(0.001)	(0.652)	(0.391)	(0.001)			
7. *Tenure*	0.112	0.122	0.015	−0.169	−0.201	0.024	1.000	
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
8. *FirmAge*	0.093	0.115	0.010	−0.291	−0.428	0.025	0.337	1.000
	(0.001)	(0.001)	(0.026)	(0.001)	(0.001)	(0.001)	(0.001)	

Pearson correlation matrix for the full sample. *CAR* is cumulative market-adjusted returns for the 12-month period ending three months after the fiscal year-end, where market-adjusted returns are the difference between raw returns and value-weighted CRSP market returns. *E* is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. *Stock Rankings* and *Debt Ratings* represent numerical values of S&P common stock rankings and S&P senior debt ratings, respectively. *Stock Rankings* (*Debt Ratings*) are assigned a value of 1 if a firm's S&P common stock (senior debt) is rated as A+ (AAA), and as the S&P common stock rankings (debt ratings) decline, the numerical value increases by 1. *FEPS* is the mean annual one-year-ahead earnings per share forecasts for year t issued immediately following the earnings announcement for year $t-1$. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. *FirmAge* measures the number of years that the firm has been publicly traded as of the fiscal year-end.
The p-values for correlation coefficients are reported in parentheses.

regression without the control variables using the full sample. More important, $\beta_3 + \beta_4$ is positive and significant (0.012, t-statistic = 3.45). Regression estimates indicate that investors on average pay a premium of 2.4 percent (= 0.012/0.507) for earnings as the auditor-client relationship increases by an additional year. The coefficient on *Tenure* (β_5) is also significant (0.005, t-statistic = 4.06), indicating that auditor tenure is positively associated with returns.

The ERC is 0.432 (t-statistic = 9.00) in the second regression when we include a number of control variables. $\beta_3 + \beta_4$ remains positive and significant (0.008, t-statistic = 2.48). The parameter estimates suggest that, after controlling for the other determinants of ERC, investors pay a premium of 2 percent (= 0.008/0.432) for earnings as the audit engagement increases by an additional year. Even though the magnitude of $\beta_3 + \beta_4$ is smaller in the second regression, we find it reassuring that the results continue to be significant even after including many control variables. When we estimate the augmented model using the restricted sample for which the auditor-client relationship lasts for at least five years, we find similar results. Specifically, $\beta_3 + \beta_4$ is positive and significant at the 5 percent level (0.008, t-statistic = 2.03) in the third regression.

The parameter estimates of the components of ERC provide added insights into investor perceptions of the time-series properties of earnings. The coefficients on *E* and ΔE in all three regressions are positive and significant at the 1 percent level. Considering that earnings levels (earnings changes) are more important for pricing decisions when earnings contain transitory (permanent) components (Easton and Harris 1991; Ali and Zarowin 1992), the results suggest that investors view reported earnings as containing both transitory and permanent components. Individual coefficients on *E*Tenure* (β_3) and Δ*E*Tenure* (β_4) also

TABLE 3
Earnings Response Coefficients and Perceptions of Investors

		Full Sample		
Variables	**(Coefficients)**	**(1)**	**(2)**	**Restricted Sample**
Intercept	(α)	−0.037 (−5.82)***	−0.096 (−6.40)***	−0.088 (−5.87)***
E	(β_1)	0.356 (6.91)***	0.261 (4.51)***	0.195 (2.92)***
ΔE	(β_2)	0.151 (8.09)***	0.171 (5.04)***	0.232 (5.60)***
	$(\beta_1 + \beta_2)$	0.507 (10.30)***	0.432 (9.00)***	0.427 (7.09)***
*E*Tenure*	(β_3)	0.007 (2.49)**	0.006 (1.80)*	0.009 (2.18)**
Δ*E*Tenure*	(β_4)	0.005 (3.39)***	0.002 (1.34)	−0.001 (−0.78)
	$(\beta_3 + \beta_4)$	0.012 (3.45)***	0.008 (2.48)**	0.008 (2.03)**
Tenure	(β_5)	0.005 (4.06)***	0.003 (5.68)***	0.003 (5.29)***
Control Variables				
*E*FirmAge* (β_6)/Δ*E*FirmAge* (β_7)				
	$(\beta_6 + \beta_7)$		−0.002 (−1.75)	−0.002 (−1.71)
*E*Big4* (β_8)/Δ*E*Big4* (β_9)				
	$(\beta_8 + \beta_9)$		−0.091 (−2.25)**	−0.117 (−2.11)**
*E*Growth* (β_{10})/Δ*E*Growth* (β_{11})				
	$(\beta_{10} + \beta_{11})$		0.023 (1.35)	0.035 (1.92)*
*E*Persistence* (β_{12})/Δ*E*Persistence* (β_{13})				
	$(\beta_{12} + \beta_{13})$		0.084 (2.42)**	0.079 (2.19)**
*E*Volatility* (β_{14})/Δ*E*Volatility* (β_{15})				
	$(\beta_{14} + \beta_{15})$		−0.083 (−7.40)***	−0.092 (−7.52)***
*E*Beta* (β_{16})/Δ*E*Beta* (β_{17})				
	$(\beta_{16} + \beta_{17})$		−0.061 (−1.87)*	−0.070 (−1.99)*
*E*Size* (β_{18})/Δ*E*Size* (β_{19})				
	$(\beta_{18} + \beta_{19})$		0.119 (9.27)***	0.132 (8.79)***
*E*Leverage* (β_{20})/Δ*E*Leverage* (β_{21})				
	$(\beta_{20} + \beta_{21})$		−0.290 (−4.74)***	−0.296 (−4.21)***
*E*Regulation* (β_{22})/Δ*E*Regulation* (β_{23})				
	$(\beta_{22} + \beta_{23})$		−0.061 (−1.38)	−0.051 (−1.27)
FirmAge	(β_{24})		0.001 (4.58)***	0.001 (4.47)***
Big4	(β_{25})		0.047 (4.00)***	0.041 (3.43)***
Growth	(β_{26})		0.023 (5.41)***	0.024 (5.60)***
Persistence	(β_{27})		0.014 (1.93)*	0.015 (2.28)**
Volatility	(β_{28})		−0.025 (−7.75)***	−0.025 (−8.53)***
Beta	(β_{29})		−0.005 (−0.18)	−0.003 (−0.10)
Size	(β_{30})		−0.005 (−1.23)	−0.005 (−1.39)
Leverage	(β_{31})		−0.018 (−0.71)	−0.014 (−0.53)
Regulation	(β_{32})		0.045 (4.01)***	0.044 (3.72)***
Yearly Observations		3,044–4,122	3,044–4,122	2,897–3,803
Adjusted R^2		0.039–0.162	0.111–0.372	0.109–0.369

(*continued on next page*)

TABLE 3 (continued)

---

*, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 level, respectively, for a two-tailed test. The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. The dependent variable is cumulative market-adjusted returns for the 12-month period ending three months after the fiscal year-end (*CAR*). Market-adjusted returns are the difference between raw returns and value-weighted CRSP market returns. E is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. We suppress the individual coefficients on E*control variable and ΔE*control variable; instead, we report the sum of the two coefficients. The control variables are defined as follows: *FirmAge* measures the number of years that the firm is publicly traded as of the fiscal year-end; *Big4* is an indicator variable that equals 1 when the client's auditor is a large accounting firm; *Growth* is the sum of the market value of equity and the book value of debt scaled by the book value of total assets; *Persistence* (*Volatility*) is the first-order autocorrelation (standard deviation) of income before extraordinary items per share for the past 16 quarters; *Beta* is the systematic risk computed using the past 60 monthly stock returns; *Size* is the logarithmic transformation of the fiscal year-end market value of equity of the prior year; *Leverage* is the ratio of total debt to total assets; and *Regulation* is an indicator variable that equals 1 for firms in a regulated industry with two-digit SIC codes between 40 and 49 or between 60 and 63, otherwise it equals 0.
The reported coefficients are the average of yearly coefficients from 1990 to 2000, and the corresponding t-statistics in parentheses are based on the distribution of the yearly coefficients.

provide some insights: (1) β_3 is larger than β_4 in all three regressions; and (2) β_3 is statistically significant in all three regressions at the 10 percent level or below, while β_4 is significant in only one regression. The incremental importance of earnings levels in pricing decisions suggests that, as the auditor-client relationship lengthens, investors perceive reported earnings as being more transitory.

We briefly discuss the results of the control variables because they have been extensively analyzed in prior studies (Subramanyam and Wild 1996; Warfield et al. 1995; Dhaliwal and Reynolds 1994; Collins and Kothari 1989). To conserve space, we do not report the individual coefficients on the interactions between each control variable and $E/\Delta E$. Instead, we report the sum of the two interaction coefficients. Controlling for other determinants of ERC, our results based on the restricted sample suggest that the ERC does not vary with *FirmAge* ($\beta_6 + \beta_7$ is insignificant). Since $\beta_8 + \beta_9$ is negative and significant, our results indicate that the ERC is lower for client firms of large auditing firms (*Big4*), which is inconsistent with the findings of Teoh and Wong (1993). As in prior studies, we find that the ERC is positively associated with *Growth*, *Persistence*, and *Size*, and is negatively associated with *Volatility*, *Beta*, *Leverage*, and *Regulation*. Also, the coefficients on *FirmAge, Big4*, *Growth*, *Persistence*, *Volatility*, and *Regulation* are all significant at the 5 percent level or below.

Bias in the Estimated ERC from Prices Leading Earnings

A potential concern in Table 3 is that the estimated ERC from a contemporaneous returns-earnings regression might be biased downward if the market anticipates current year earnings more than one year prior to the earnings release (Kothari 1992; Kothari and Sloan 1992). The estimated ERC could be biased from using a 12-month contemporaneous period to compute returns or from using market value of equity as a deflator.[12] Further, the downward bias in the estimated ERC is expected to be larger for firms with longer tenure. Given that prior studies generally find that earnings quality improves with extended tenure, and considering that high-quality earnings are easier to predict, the market is more likely to anticipate current-year earnings for firms with longer tenure.

We mitigate potential biases in the estimated ERC in two ways. First, as in DeFond and Hung (2003) and Collins and Kothari (1989), we measure returns over a 15-month period ending three months after the fiscal year-end. Increasing the returns measurement window enhances the probability of including the period in which the market anticipates current earnings thereby reducing any measurement error. Second, as in Subramanyam (1996), we deflate earnings by total assets at the beginning of the year. Since total assets are typically based on historical cost rather than market value, any bias in the estimated coefficients from prices leading earnings is likely to be reduced from using total assets as a deflator.

Panel A of Table 4 reports the results from using cumulative market-adjusted returns measured over 15 months ($CAR_{15\text{-}month}$) as the dependent variable. While the magnitude of ERC ($\beta_1 + \beta_2$) is smaller, we find that the size of the sum of the coefficients on *E*Tenure* and Δ*E*Tenure* ($\beta_3 + \beta_4$) is 25 to 38 percent larger than that reported in Table 3. One reason why the ERC is smaller may be because of the large number of control variables. When we estimate the regressions in Panel A without the control variables, we find that the estimated ERC becomes larger. Specifically, unreported results show that the ERC for the full sample without the control variables is 0.568, while the corresponding estimate is 0.507 in Table 3.

We find similar results using total assets as a deflator for E and ΔE. The estimate of $\beta_3 + \beta_4$ using the full (restricted) sample is 0.027 (0.025) in Panel B of Table 4, while the corresponding estimate in Table 3 is 0.008 (0.008). Thus, the magnitude of $\beta_3 + \beta_4$ is 213 to 238 percent larger in Panel B compared to that reported in Table 3. The magnitude of ERC is also larger in Panel B than that in Table 3. Hence, the results in Table 4 suggest that investors are more likely to anticipate current-year earnings more than one year ahead of the earnings release for firms with extended auditor-client relationships.

Overall, the results from Tables 3 and 4 are consistent with the hypothesis that investors perceive earnings quality as improving with longer auditor tenure. Our results suggest that investors view audit quality as improving with auditor tenure.

Perceptions of Information Intermediaries and Auditor Tenure

S&P Stock Rankings and Debt Ratings

Unlike stock returns, stock rankings and debt ratings are possibly correlated over time. Therefore, the residuals from yearly regressions may not be independent, which could lead to biased t-statistics. We compute an autocorrelation-corrected t-statistic (t-statistic^{AC}) by comparing the mean coefficients with the time-series-based standard error that accounts for

---

[12] If prices lead earnings, current period earnings contain information that is "new" to the market and information that is "stale" or "old." Given that stock prices change in response to new information, the independent variable contains errors of measurement from using current period earnings as the regressor (also known as errors-in-variable problem). Measurement error on the independent variable biases the slope coefficient downward because of its correlation with the regression disturbance term (Greene 1993). Similarly, deflating earnings by market value of equity might inject a downward bias in the estimated coefficients when the market anticipates current period earnings at the beginning of the year.

TABLE 4
Earnings Response Coefficients and Perceptions of Investors: Sensitivity Analysis

Variables	(Coefficients)	Full Sample	Restricted Sample
Panel A: $CAR_{15\text{-month}}$ is the Dependent Variable			
E	(β_1)	0.217 (2.62)**	0.165 (1.95)*
ΔE	(β_2)	0.169 (3.80)***	0.248 (4.07)***
	$(\beta_1 + \beta_2)$	0.386 (5.50)***	0.413 (5.21)***
*E*Tenure*	(β_3)	0.008 (1.65)	0.011 (1.92)*
ΔE**Tenure*	(β_4)	0.003 (1.30)	−0.001 (−0.91)
	$(\beta_3 + \beta_4)$	0.011 (2.84)***	0.010 (2.24)**
Tenure	(β_5)	0.004 (5.99)***	0.003 (5.37)***
Control Variables (as in Table 3)		Included	Included
Yearly Observations		3,044–4,122	2,897–3,803
Adjusted R^2		0.145–0.322	0.154–0.327
Panel B: Earnings are Scaled by Total Assets			
E	(β_1)	0.415 (5.06)***	0.393 (5.09)***
ΔE	(β_2)	0.379 (3.99)***	0.468 (3.69)***
	$(\beta_1 + \beta_2)$	0.794 (8.33)***	0.861 (8.57)***
*E*Tenure*	(β_3)	0.006 (0.84)	0.012 (1.16)
ΔE**Tenure*	(β_4)	0.021 (2.46)**	0.013 (1.58)
	$(\beta_3 + \beta_4)$	0.027 (2.41)**	0.025 (2.38)**
Tenure	(β_5)	0.002 (2.68)**	0.001 (0.94)
Control Variables (as in Table 3)		Included	Included
Yearly Observations		3,218–4,207	3,063–3,888
Adjusted R^2		0.095–0.278	0.091–0.288

*, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 level, respectively, for a two-tailed test. The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. In Panel A, the dependent variable is cumulative market-adjusted returns for the 15-month period ending three months after the fiscal year-end ($CAR_{15\text{-month}}$). Market-adjusted returns are the difference between raw returns and value-weighted CRSP market returns. *E* is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. In Panel B, the dependent variable is *CAR* measured over the 12-month period ending three months after the fiscal year-end. *E* is income before extraordinary items deflated by total assets at the beginning of the year. ΔE is the difference between income before extraordinary items for the current year and that of last year deflated by total assets at the beginning of the year. The results of all the control variables in Panels A and B are not reported for brevity. The reported coefficients are the average of yearly coefficients from 1990 to 2000, and the corresponding t-statistics in parentheses are based on the distribution of the yearly coefficients.

any dependence in the estimated coefficients, as in Abarbanell and Bernard (2000).[13] Since our focus is on auditor tenure, we compute t-statistic^{AC} for coefficients that involve *Tenure* only.

---

[13] We assume that serial correlation is second-order autoregressive, so that the correction involves multiplying the yearly standard errors by the factor $\sqrt{\left[\frac{(1+r)}{(1-r)}\right] - \left[\frac{2*r*(1-r^n)}{n*(1-r)^2}\right]}$ where *n* is the number of observations used to compute the mean coefficient and *r* is the average of the first- and second-order autocorrelations estimated from the series of annual coefficients. The correction factor is not applied when estimated autocorrelation is less than zero.

Table 5 reports the regression results of common stock rankings on auditor tenure. The first column in Table 5 presents the results from the reduced model using the full sample.

TABLE 5
Stock Rankings and Perceptions of Information Intermediaries

Variables	**(Coefficients)**	**Full Sample**		
		(1)	**(2)**	**Restricted Sample**
Intercept	(α)	5.746 (100.31)***	6.610 (140.79)***	6.614 (149.22)***
_E_	(β_1)	−1.969 (−5.25)***	−0.208 (−0.91)	−0.366 (−1.01)
_ΔE_	(β_2)	1.314 (4.77)***	−0.148 (−0.58)	−0.030 (−0.08)
	$(\beta_1 + \beta_2)$	−0.655 (−5.25)***	−0.356 (−2.03)**	−0.396 (−1.84)*
E*Tenure	(β_3)	−0.100 (−3.80)***	−0.061 (−5.16)***	−0.063 (−5.87)***
ΔE*Tenure	(β_4)	0.046 (2.39)**	0.021 (1.93)*	0.023 (1.96)*
	$(\beta_3 + \beta_4)$	−0.054 (−3.76)***	−0.040 (−3.31)***	−0.040 (−2.97)***
Tenure	(β_5)	−0.068 (−6.05)***	−0.023 (−9.33)***	−0.024 (−9.79)***
Control Variables				
E*FirmAge (β_6)/_ΔE*FirmAge_ (β_7)				
	$(\beta_6 + \beta_7)$		0.017 (2.45)**	0.021 (2.72)**
E*Big4 (β_8)/_ΔE*Big4_ (β_9)				
	$(\beta_8 + \beta_9)$		0.264 (2.54)**	0.262 (2.61)**
E*Growth (β_{10})/_ΔE*Growth_ (β_{11})				
	$(\beta_{10} + \beta_{11})$		−0.134 (−1.66)	−0.158 (−1.77)
E*Persistence (β_{12})/_ΔE*Persistence_ (β_{13})				
	$(\beta_{12} + \beta_{13})$		0.536 (3.40)***	0.625 (4.23)***
E*Volatility (β_{14})/_ΔE*Volatility_ (β_{15})				
	$(\beta_{14} + \beta_{15})$		0.189 (2.60)**	0.193 (2.64)**
E*Beta (β_{16})/_ΔE*Beta_ (β_{17})				
	$(\beta_{16} + \beta_{17})$		0.265 (4.21)***	0.320 (3.79)***
E*Size (β_{18})/_ΔE*Size_ (β_{19})				
	$(\beta_{18} + \beta_{19})$		−0.208 (−4.40)***	−0.226 (−4.09)***
E*Leverage (β_{20})/_ΔE*Leverage_ (β_{21})				
	$(\beta_{20} + \beta_{21})$		0.489 (2.62)**	0.438 (1.83)*
E*Regulation (β_{22})/_ΔE*Regulation_ (β_{23})				
	$(\beta_{22} + \beta_{23})$		0.038 (0.28)	0.040 (0.24)
FirmAge	(β_{24})		0.001 (0.88)	0.000 (0.46)
Big4	(β_{25})		0.250 (11.51)***	0.256 (14.51)***
Growth	(β_{26})		0.033 (2.66)**	0.030 (3.19)***
Persistence	(β_{27})		−0.032 (−0.79)	−0.023 (−0.53)
Volatility	(β_{28})		0.309 (15.93)***	0.329 (15.88)***
Beta	(β_{29})		0.372 (8.71)***	0.373 (9.44)***
Size	(β_{30})		−0.375 (−26.10)***	−0.371 (−24.15)***
Leverage	(β_{31})		0.407 (5.77)***	0.396 (5.63)***
Regulation	(β_{32})		−0.207 (−6.00)***	−0.205 (−6.10)***
Yearly Observations		2,019–2,637	2,019–2,637	1,949–2,487
Adjusted R^2		0.125–0.169	0.463–0.523	0.448–0.519

(continued on next page)

TABLE 5 (continued)

---

*, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 level, respectively, for a two-tailed test.
The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. The dependent variable is S&P common stock rankings converted into numerical values (*Stock Rankings*). *Stock Rankings* take the values of 1 to 7 representing S&P common stock rankings of A+, A, A−, B+, B, B−, and C, respectively. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. E is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is defined as the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. We suppress the individual coefficients on E*control variable and ΔE*control variable; instead, we report the sum of the two coefficients. The control variables are defined as follows: *FirmAge* measures the number of years that the firm is publicly traded as of the fiscal year-end; *Big4* is an indicator variable that equals 1 when the client's auditor is a large accounting firm; *Growth* is the sum of the market value of equity and the book value of debt scaled by the book value of total assets; *Persistence* (*Volatility*) is the first-order autocorrelation (standard deviation) of income before extraordinary items per share for the past 16 quarters; *Beta* is the systematic risk computed using the past 60 monthly stock returns; *Size* is the logarithmic transformation of the fiscal year-end market value of equity of the prior year; *Leverage* is the ratio of total debt to total assets; and *Regulation* is an indicator variable that equals 1 for firms in a regulated industry with two-digit SIC codes between 40 and 49 or between 60 and 63, and 0 otherwise.
The reported coefficients are the average of yearly coefficients from 1990 to 2000, and the corresponding t-statistics in parentheses are based on the distribution of the yearly coefficients.

As in Table 3, ERC, a proxy for earnings quality as perceived by ratings agencies, is defined as the sum of the coefficients on E and ΔE.[14] The ERC is negative (−0.655) and statistically significant at the 1 percent level (t-statistic = −5.25). $\beta_3 + \beta_4$, which measures the joint effects of tenure and earnings on rankings, is also negative (−0.054) and significant at the 1 percent level (t-statistic = −3.76; t-statistic^{AC} = −3.14). All else equal, for a given increase in earnings, *Stock Rankings* improve by an additional 8.2 percent (= −0.054/ −0.655) when the audit engagement increases by an added year.

The coefficient on *Tenure* (β_5) is negative (−0.068) and significant at the 1 percent level (t-statistic = −6.05; t-statistic^{AC} = −3.52). Other things remaining constant, if auditor tenure for a firm is one standard deviation above the mean value of *Tenure*, the stock ranking is lower by 0.327 (= −0.068 × 4.807), compared to the average stock ranking of 5.096 for the sample used in the regression.

The second column in Table 5 reports the results from the augmented model using the full sample. $\beta_3 + \beta_4$ remains negative (−0.040) and significant (t-statistic = −3.31; t-statistic^{AC} = −3.31). β_5 is also negative (−0.023) and significant (t-statistic = −9.33; t-statistic^{AC} = −6.59). The results are very similar when we use the restricted sample (reported in the last column). $\beta_3 + \beta_4$ is negative (−0.040) and significant (t-statistic = −2.97; t-statistic^{AC} = −2.97), and β_5 is highly significant (−0.024, t-statistic = −9.79; t-statistic^{AC} = −6.78). Our results suggest that rating agencies place a greater weight on reported earnings as auditor tenure increases. Given that S&P incorporates the quality of accounting information as a factor in establishing its rankings/ratings (Standard & Poor's 1982), the importance of tenure suggests that, as tenure increases, independent rating agencies perceive audited financial statements as being more reliable.

We briefly discuss how various control variables affect *Stock Rankings* (based on the full sample). The coefficients on the interactions between earnings and *FirmAge*, *Persistence*, *Volatility*, *Beta*, and *Leverage* are positive and significant at the 5 percent level or below. These results suggest that the effect of earnings on *Stock Rankings* is smaller for mature/risky firms. The positive sign of the coefficient on the interactions between earnings and *Big4* ($\beta_8 + \beta_9$) is puzzling. The coefficient on the interactions between earnings and *Size* ($\beta_{18} + \beta_{19}$) is negative and significant at the 1 percent level, indicating that the effect of earnings on *Stock Rankings* is larger for bigger firms. The coefficients on the interactions between earnings and *Growth* ($\beta_{10} + \beta_{11}$) and those between earnings and *Regulation* ($\beta_{22} + \beta_{23}$) are both insignificant.

---

[14] Although the term ERC or earnings response coefficient is typically used for earnings coefficients in returns-earnings regressions, we use the term in a broader context to refer to earnings coefficients for the returns, stock rankings, debt ratings, and analyst forecasts regressions.

The control variables are also included as separate independent variables. The positive sign of the coefficient on *Big4* is again puzzling. The coefficients on *Volatility*, *Beta*, and *Leverage* are all positive and significant, which suggests that higher risk has an unfavorable impact on stock rankings. The coefficient on *Growth* is also positive, which is inconsistent with our expectations. If our model does not adequately control for risk, growth firms are more likely to be associated with unfavorable *Stock Rankings* because higher growth could proxy for larger risk. The coefficients on *Size* and *Regulation* are negative, which suggests that stock rankings are more favorable for larger firms and those in regulated industry. The coefficients on *FirmAge* and *Persistence* are insignificant.

Table 6 presents the regression results of the association between debt ratings and auditor tenure. The first column in Table 6 presents the results of estimating the reduced model using the full sample. Consistent with prior research, the ERC is negative and significant (−4.568, t-statistic = −5.87). Thus, earnings have a favorable influence on debt ratings. The sum of the coefficients on *E*Tenure* and *ΔE*Tenure* ($\beta_3 + \beta_4$) is also negative, but insignificant (−0.093, t-statistic = −1.03; t-statistic^{AC} = −0.98). In contrast, the coefficient on *Tenure* (β_5) is negative and significant (−0.183, t-statistic = −6.33; t-statistic^{AC} = −3.42), consistent with the findings of Mansi et al. (2004).

Once we include the control variables in Table 6, the ERC is insignificant in either the second or the third regression. Although $\beta_3 + \beta_4$ is marginally significant (−0.216, t-statistic = −1.98; t-statistic^{AC} = −1.85) for the full sample, it is insignificant (−0.240, t-statistic = −1.66; t-statistic^{AC} = −1.62) for the restricted sample. On the other hand, β_5 remains negative and significant for the full sample (−0.065, t-statistic = −4.78; t-statistic^{AC} = −2.70) and the restricted sample (−0.067, t-statistic = −4.51; t-statistic^{AC} = −2.70). Thus, although longer tenure is associated with improved debt ratings, the association between earnings and debt ratings does not appear to vary with tenure.[15]

In a recent study, Mansi et al. (2004) find that auditor tenure is negatively related to the cost of debt and that the relation is more pronounced for non-investment-grade debt. This finding suggests that auditor tenure could influence debt ratings differently for investment and non-investment-grade debt. Following Mansi et al. (2004), we separately estimate the augmented model for investment and non-investment-grade debt to investigate whether the impact of auditor tenure on debt ratings depends on the quality of debt. When we partition our sample into investment and non-investment-grade debt subsamples, the results are very similar to those reported in Table 6. For example, $\beta_3 + \beta_4$ is negative but not significant in either subsample, but β_5 is negative and significant for both subsamples (for brevity, the results are not tabulated).

In sum, the regression results from stock rankings and debt ratings provide modest evidence that independent rating agencies perceive firms with longer auditor tenure as providing more reliable financial information.

---

[15] Since the same control variables appear in the stock rankings and debt ratings regressions, and the results are similar across the two sets of regressions, we do not repeat our discussion of the control variables for the debt ratings regressions.

Financial Analysts' Earnings Forecasts

Table 7 reports the results of how auditor tenure affects the association between one-year-ahead consensus earnings per share forecasts and earnings per share reported just prior to the forecasts. As in the previous tables, our interest is in the sum of the coefficients on EPS_{t-1}* *Tenure* and ΔEPS_{t-1}**Tenure* ($\beta_3 + \beta_4$). The first column in Table 7 reports the

TABLE 6
Debt Ratings and Perceptions of Information Intermediaries

		Full Sample		
Variables	**(Coefficients)**	**(1)**	**(2)**	**Restricted Sample**
Intercept	(α)	11.101 (64.35)***	14.387 (42.74)***	14.032 (31.54)***
E	(β_1)	−8.796 (−3.37)***	−1.378 (−0.21)	2.596 (0.40)
ΔE	(β_2)	4.228 (2.13)**	14.567 (1.13)	11.871 (0.88)
	($\beta_1 + \beta_2$)	−4.568 (−5.87)***	13.189 (1.02)	14.467 (1.13)
*E*Tenure*	(β_3)	−0.169 (−0.84)	0.014 (0.18)	0.020 (0.22)
ΔE**Tenure*	(β_4)	0.076 (0.54)	−0.230 (−2.06)**	−0.260 (−1.89)*
	($\beta_3 + \beta_4$)	−0.093 (−1.03)	−0.216 (−1.98)*	−0.240 (−1.66)
Tenure	(β_5)	−0.183 (−6.33)***	−0.065 (−4.78)***	−0.067 (−4.51)***
Control Variables				
*E*FirmAge* (β_6)/ΔE**FirmAge* (β_7)				
	($\beta_6 + \beta_7$)		0.040 (3.11)***	0.040 (2.63)**
*E*Big4* (β_8)/ΔE**Big4* (β_9)				
	($\beta_8 + \beta_9$)		−13.630 (−1.05)	−15.302 (−1.17)
*E*Growth* (β_{10})/ΔE**Growth* (β_{11})				
	($\beta_{10} + \beta_{11}$)		1.333 (1.61)	1.641 (1.54)
*E*Persistence* (β_{12})/ΔE**Persistence* (β_{13})				
	($\beta_{12} + \beta_{13}$)		1.284 (3.17)***	1.280 (2.97)***
*E*Volatility* (β_{14})/ΔE**Volatility* (β_{15})				
	($\beta_{14} + \beta_{15}$)		0.068 (0.57)	0.114 (0.67)
*E*Beta* (β_{16})/ΔE**Beta* (β_{17})				
	($\beta_{16} + \beta_{17}$)		1.011 (2.11)**	1.083 (2.26)**
*E*Size* (β_{18})/ΔE**Size* (β_{19})				
	($\beta_{18} + \beta_{19}$)		−0.594 (−2.67)**	−0.552 (−2.25)**
*E*Leverage* (β_{20})/ΔE**Leverage* (β_{21})				
	($\beta_{20} + \beta_{21}$)		0.449 (0.29)	0.512 (0.32)
*E*Regulation* (β_{22})/ΔE**Regulation* (β_{23})				
	($\beta_{22} + \beta_{23}$)		−1.127 (−1.84)*	−1.271 (−1.99)*
FirmAge	(β_{24})		−0.014 (−7.46)***	−0.014 (−7.13)***
Big4	(β_{25})		1.473 (3.13)***	1.914 (3.80)***
Growth	(β_{26})		−0.102 (−1.94)*	−0.108 (−1.98)*
Persistence	(β_{27})		0.306 (4.78)***	0.305 (4.60)***
Volatility	(β_{28})		0.254 (7.55)***	0.223 (5.36)***
Beta	(β_{29})		1.359 (12.12)***	1.333 (11.66)***
Size	(β_{30})		−1.087 (−31.65)***	−1.088 (−29.81)***
Leverage	(β_{31})		3.676 (13.00)***	3.663 (12.30)***
Regulation	(β_{32})		−1.025 (−10.18)***	−1.017 (−10.20)***
Yearly Observations		603–1,124	603–1,124	598–1,051
Adjusted R^2		0.132–0.245	0.602–0.699	0.596–0.697

(*continued on next page*)

TABLE 6 (continued)

---

*, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 level, respectively, for a two-tailed test.
The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. The dependent variable is S&P senior debt ratings converted into numerical values (*Debt Ratings*). *Debt Ratings* are assigned a value of 1 if a firm's S&P senior debt is rated as AAA. As the S&P debt ratings decline, the numerical value increases by 1. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. E is income before extraordinary items deflated by market value of equity at the beginning of the year. ΔE is defined as the difference between income before extraordinary items for the current year and that of last year deflated by market value of equity at the beginning of the year. We suppress the individual coefficients on E*control variable and ΔE*control variable; instead, we report the sum of the two coefficients. The control variables are defined as follows: *FirmAge* measures the number of years that the firm is publicly traded as of the fiscal year-end; *Big4* is an indicator variable that equals 1 when the client's auditor is a large accounting firm; *Growth* is the sum of the market value of equity and the book value of debt scaled by the book value of total assets; *Persistence* (*Volatility*) is the first-order autocorrelation (standard deviation) of income before extraordinary items per share for the past 16 quarters; *Beta* is the systematic risk computed using the past 60 monthly stock returns; *Size* is the logarithmic transformation of the fiscal year-end market value of equity of the prior year; *Leverage* is the ratio of total debt to total assets; and *Regulation* is an indicator variable that equals 1 for firms in a regulated industry with two-digit SIC codes between 40 and 49 or between 60 and 63, and 0 otherwise.
The reported coefficients are the average of yearly coefficients from 1990 to 2000, and the corresponding t-statistics in parentheses are based on the distribution of the yearly coefficients.

results of the reduced model for the full sample. We find that reported earnings are positively associated with analysts' earnings forecasts. The ERC ($\beta_1 + \beta_2$) is positive and highly significant (1.008, t-statistic = 20.29). Moreover, $\beta_3 + \beta_4$ is also positive and significant (0.019, t-statistic = 2.68). Other things remaining constant, our results suggest that analysts appear to attach greater importance to the most recent reported earnings in forming their expectations about future earnings as auditor tenure grows longer.

When we estimate the augmented model that includes a number of control variables, $\beta_3 + \beta_4$ remains positive and significant (0.016, t-statistic = 2.90) in the second regression. The results are very similar when we estimate the augmented model using the restricted sample. $\beta_3 + \beta_4$ continues to be positive and significant in the third regression (0.016, t-statistic = 2.51). On the other hand, β_5 is insignificant in all three regressions.

The individual parameter estimates of ERC and $\beta_3 + \beta_4$ provide some added insights. The coefficients on EPS_{t-1} (β_1) and ΔEPS_{t-1} (β_2) are both significant in the first regression. However, in the second and third regressions that include the other determinants of earnings forecasts, β_2 becomes insignificant. Large earnings surprises are more likely to be temporary and, therefore, potentially less useful to analysts in predicting future earnings (Barron et al. 2002; Stickel 1989). In contrast, the coefficient on ΔEPS_{t-1}**Tenure* (β_4) is significant in all three regressions at the 10 percent level or below. The results suggest that although past earnings changes do not influence earnings forecasts, they are perceived as more relevant for earnings forecasts as the auditor-client relationship increases.

Based on the results from the restricted sample, we briefly discuss how the various control variables affect analysts' earnings forecasts ($FEPS_t$). The coefficients on the interactions between reported earnings and *FirmAge*, *Leverage*, and *Analysts* are all negative and significant at the 10 percent level or below. These results suggest that the effect of reported earnings on $FEPS_t$ is smaller for mature/risky firms and for firms with larger analyst following. The interaction between reported earnings and *Size* is positive and significant, which suggests that the impact of reported earnings on analysts' earnings forecasts is greater for bigger firms. Also, the individual coefficients on *FirmAge*, *Volatility*, *Leverage*, and *Analysts* are all positive and significant at the 5 percent level or below, which suggests that analysts' earnings forecasts are higher for mature firms, for firms with higher risk, and those with larger analyst following. The coefficient on *Size* is negative, which suggests that earnings per share forecasts are smaller for bigger firms. One explanation is that large, politically sensitive firms are more likely to report lower earnings to reduce their political costs and, therefore, analysts' earnings forecasts are also lower for larger firms.

TABLE 7
Analyst Earnings Forecasts and Perceptions of Information Intermediaries

		Full Sample		
Variables	**(Coefficients)**	**(1)**	**(2)**	**Restricted Sample**
Intercept	(α)	0.217 (8.84)***	0.366 (15.47)***	0.364 (14.87)***
EPS_{t-1}	(β_1)	0.834 (22.34)***	0.532 (15.28)***	0.530 (15.64)***
ΔEPS_{t-1}	(β_2)	0.174 (4.59)***	−0.068 (−0.66)	−0.074 (−0.70)
	($\beta_1 + \beta_2$)	1.008 (20.29)***	0.464 (4.62)***	0.456 (4.49)***
$EPS_{t-1}*Tenure$	(β_3)	0.005 (1.75)	0.005 (1.79)	0.005 (1.54)
$\Delta EPS_{t-1}*Tenure$	(β_4)	0.014 (2.00)**	0.011 (2.11)**	0.011 (1.95)*
	($\beta_3 + \beta_4$)	0.019 (2.68)**	0.016 (2.90)***	0.016 (2.51)**
Tenure	(β_5)	−0.000 (−0.15)	−0.003 (−1.16)	−0.002 (−0.76)
Control Variables				
$EPS_{t-1}*FirmAge$ (β_6)/$\Delta EPS_{t-1}*FirmAge$ (β_7)				
	($\beta_6 + \beta_7$)		−0.004 (−2.83)***	−0.004 (−2.90)***
$EPS_{t-1}*Big4$ (β_8)/$\Delta EPS_{t-1}*Big4$ (β_9)				
	($\beta_8 + \beta_9$)		−0.060 (−1.29)	−0.077 (−1.49)
$EPS_{t-1}*Growth$ (β_{10})/$\Delta EPS_{t-1}*Growth$ (β_{11})				
	($\beta_{10} + \beta_{11}$)		−0.008 (−0.33)	−0.010 (−0.41)
$EPS_{t-1}*Volatility$ (β_{12})/$\Delta EPS_{t-1}*Volatility$ (β_{13})				
	($\beta_{12} + \beta_{13}$)		−0.032 (−0.67)	−0.033 (−0.67)
$EPS_{t-1}*Beta$ (β_{14})/$\Delta EPS_{t-1}*Beta$ (β_{15})				
	($\beta_{14} + \beta_{15}$)		−0.011 (−0.37)	−0.012 (−0.38)
$EPS_{t-1}*Size$ (β_{16})/$\Delta EPS_{t-1}*Size$ (β_{17})				
	($\beta_{16} + \beta_{17}$)		0.140 (7.17)***	0.147 (8.00)***
$EPS_{t-1}*Leverage$ (β_{18})/$\Delta EPS_{t-1}*Leverage$ (β_{19})				
	($\beta_{18} + \beta_{19}$)		−0.235 (−2.18)**	−0.225 (−1.99)*
$EPS_{t-1}*Regulation$ (β_{20})/$\Delta EPS_{t-1}*Regulation$ (β_{21})				
	($\beta_{20} + \beta_{21}$)		−0.041 (−1.26)	−0.034 (−1.04)
$EPS_{t-1}*Analysts$ (β_{22})/$\Delta EPS_{t-1}*Analysts$ (β_{23})				
	($\beta_{22} + \beta_{23}$)		−0.013 (−3.85)***	−0.014 (−4.20)***
FirmAge	(β_{24})		0.003 (4.92)***	0.004 (4.95)***
Big4	(β_{25})		0.028 (1.39)	0.034 (1.54)
Growth	(β_{26})		−0.003 (−0.91)	−0.003 (−1.00)
Volatility	(β_{27})		0.138 (3.04)***	0.138 (2.86)***
Beta	(β_{28})		−0.004 (−0.28)	−0.002 (−0.18)
Size	(β_{29})		−0.048 (−12.29)***	−0.050 (−12.99)***
Leverage	(β_{30})		0.174 (5.06)***	0.168 (4.41)***
Regulation	(β_{31})		0.020 (0.89)	0.013 (0.55)
Analysts	(β_{32})		0.002 (1.82)*	0.002 (2.17)**
Yearly Observations		1,100–1,855	1,100–1,855	1,074–1,735
Adjusted R^2		0.736–0.878	0.759–0.891	0.757–0.893

(*continued on next page*)

TABLE 7 (continued)

---

*, **, and *** denote statistical significance at the 0.10, 0.05, and 0.01 level, respectively, for a two-tailed test.
The full sample includes Compustat firms with available data from 1990 to 2000. The restricted sample consists of firms in the full sample with auditor-client relationships lasting for at least five years. The dependent variable is the mean annual one-year-ahead earnings per share forecasts for year t issued immediately following the earnings announcement for year $t-1$ ($FEPS_t$). EPS_{t-1} is annual earnings per share for year $t-1$, and ΔEPS_{t-1} is the absolute change in earnings per share for year $t-1$ defined as the difference in annual earnings per share in year $t-1$ and that in year $t-2$ ($|EPS_{t-1} - EPS_{t-2}|$). All earnings per share variables ($FEPS_t$, EPS_{t-1}, and ΔEPS_{t-1}) are obtained from I/B/E/S database. *Tenure* is the duration of the auditor-client relationship in years starting from 1982. All the control variables including *Tenure* are measured as of the end of year $t-1$. We suppress the individual coefficients on E*control variable and ΔE*control variable; instead, we report the sum of the two coefficients. The control variables are defined as follows: *FirmAge* measures the number of years that the firm is publicly traded as of the fiscal year-end; *Big4* is an indicator variable that equals 1 when the client's auditor is a large accounting firm; *Growth* is the sum of the market value of equity and the book value of debt scaled by the book value of total assets; *Volatility* is the standard deviation of income before extraordinary items per share for the past 16 quarters; *Beta* is the systematic risk computed using the past 60 monthly stock returns; *Size* is the logarithmic transformation of the fiscal year-end market value of equity of the prior year; *Leverage* is the ratio of total debt to total assets; *Regulation* is an indicator variable that equals 1 for firms in a regulated industry with two-digit SIC codes between 40 and 49 or between 60 and 63, and otherwise 0; and *Analysts* is the number of analysts providing annual earnings forecasts.
The reported coefficients are the average of yearly coefficients from 1990 to 2000, and the corresponding t-statistics in parentheses are based on the distribution of the yearly coefficients.

A key advantage of using consensus forecasts is that the aggregation process offsets individual forecast errors, or idiosyncratic errors, which results in the consensus forecasts being more accurate than the individual forecasts (Brown et al. 1985). However, one concern is that the idiosyncratic errors may not be fully diversified away when the consensus is based on a few individual forecasts (Barron et al. 1998). Although any measurement error in the dependent variable does not bias the estimated coefficients, the estimated variances are larger and therefore the t-statistics are also lower (Gujarati 1995). One solution to the potential measurement error problem is to impose a restriction on the number of individual forecasts required to compute the mean forecast. The diversification argument suggests that the measurement error declines with an increase in the number of individual forecasts.

Consistent with the diversification argument, we find that the magnitude of the t-statistics becomes larger when we impose restrictions on the number of individual forecasts. For instance, $\beta_3 + \beta_4$ is 0.01 and the associated t-statistic is 2.90 in the second regression (Table 7). However, as in Barron et al. (2002), when we require that the consensus forecasts be based on a minimum of three individual forecasts, $\beta_3 + \beta_4$ is 0.01 but the t-statistic increases to 3.40. The t-statistic jumps to 4.67 when the consensus forecast is based on a minimum of seven individual forecasts.

Thus, our results from analyst earnings forecasts are consistent with the hypothesis that financial analysts perceive earnings quality as improving with longer auditor tenure.

Sensitivity Analysis

We conduct a number of additional analyses to test the robustness of our results. Instead of tabulating the results of the additional tests, we discuss them briefly in this subsection. First, given that raw returns are frequently used to estimate contemporaneous returns-earnings models (Lundholm and Myers 2002; Warfield et al. 1995; Collins and Kothari 1989), we estimate the returns-earnings regression model using 12-month compounded raw returns (ending three months after the fiscal year-end) as the dependent variable. The results based on raw returns are very similar to those reported in Table 3 using cumulative market-adjusted returns. The sum of the coefficients on *E*Tenure* and *ΔE*Tenure* for the full and restricted samples is 0.010 (t-statistic = 2.50) and 0.009 (t-statistic = 2.15), respectively. The coefficient on *Tenure* is also positive and significant for the full sample (0.004, t-statistic = 5.52) and the restricted sample (0.003, t-statistic = 4.84).

Second, considering that cash flow might be an important determinant of debt ratings, we estimate *Debt Ratings* regressions after including operating cash flow (deflated by market value of equity at the beginning of the year) in addition to all the other variables. The inclusion of the cash flow variable does not affect the tenor of the results reported in Table 6. The coefficient on *Tenure* continues to be negative and significant for the full sample (−0.066, t-statistic = −4.91) and the restricted sample (−0.067, t-statistic = −4.54). The coefficient on cash flow is not significant.

Third, we also analyze S&P subordinated debt ratings (Compustat #320). The sample size is relatively small: the number of observations used in the yearly regressions varies between 251 and 372 for the full sample and between 245 and 358 for the restricted sample. The results from S&P subordinated debt ratings are very similar to those using senior debt ratings. The coefficient on *Tenure* is negative and significant for the full sample (−0.047, t-statistic = −2.96) and for the restricted sample (−0.031, t-statistic = −1.89). The sum of the coefficients on *E*Tenure* and *ΔE*Tenure* is negative but insignificant for both samples.

Fourth, we estimate the effects of tenure on rankings/ratings using an ordered probit model. Since *Stock Rankings* and *Debt Ratings* (the dependent variables) are discrete outcomes, conventional regression methods could yield biased estimates (Greene 1993). We find that the results in Tables 5 and 6 are not affected when we use an ordered probit model. For example, the coefficient on *Tenure* is negative and highly significant for the *Stock Rankings* regression (−0.036, $\chi^2 = 732.74$) and the *Debt ratings* regression (−0.039, $\chi^2 = 473.14$) for the full sample.

Fifth, the results reported in Table 7 are based on levels specification—earnings per share forecasts (*FEPS*) and the two reported earnings per share measures (*EPS* and *ΔEPS*) are not deflated. One concern is that the ordinary least squares estimates might be inefficient and/or biased because of scale differences (Barth and Kallapur 1996; Dechow 1994). Our conclusions remain unchanged when we estimate the analyst forecast model after (1) deflating all the earnings per share variables (*FEPS*, *EPS*, and *ΔEPS*) by the stock price per share at the beginning of the forecast period, and (2) deleting the top and bottom 1 percent of the observations for each variable. The sum of the coefficients on EPS_{t-1}**Tenure* and ΔEPS_{t-1}**Tenure* remains positive and significant for the full sample (0.015, t-statistic = 3.05) and the restricted sample (0.013, t-statistic = 2.08).

Finally, based on the specification used in prior studies, the analyst forecast model does not include earnings persistence. However, to maintain consistency with the returns, stock rankings and debt ratings specifications, we also examine whether the analyst forecasts results are robust to the inclusion of *Persistence*. When we include EPS_{t-1}**Persistence*, ΔEPS_{t-1}** Persistence*, and *Persistence* as additional independent variables, our conclusions remain unaffected. The sum of the coefficients on EPS_{t-1}**Tenure* and ΔEPS_{t-1}**Tenure* remains positive and significant for the full sample (0.017, t-statistic = 3.20) and the restricted sample (0.016, t-statistic = 2.90).

VI. CONCLUSIONS

This study provides insights into the recent debate surrounding auditor tenure, independence, and audit quality by analyzing the relationship between auditor tenure and audit quality as perceived by capital market participants. We focus on a market-based approach because academics, regulatory institutions, and practitioners frequently emphasize the importance of incorporating capital market perceptions of independence and audit quality (Dopuch et al. 2003; Ryan et al. 2001; SEC 2000; Carmichael 1999; AICPA 1994). Based on the association between perceived reliability of financial statements and auditor tenure, we infer how capital markets view tenure as affecting audit quality. All else equal, if capital market participants perceive auditor tenure as improving (impairing) audit quality, then reported financial statements are viewed as more (less) reliable for investment and credit decisions.

We analyze the perceptions of three primary users of audited financial statements: investors, independent rating agencies, and financial analysts. Using earnings response coefficients from returns-earnings regressions as a proxy for investor perceptions of earnings quality (Schipper and Vincent 2003), we find evidence consistent with the hypothesis that investors perceive earnings quality as improving with auditor tenure. Our analysis of the perceptions of independent rating agencies is based on how tenure affects the association between rankings/ratings and reported earnings. We find that the influence of reported earnings on stock rankings becomes larger with extended tenure, although the influence of reported earnings on debt ratings does not vary with tenure. Thus, our results provide limited evidence that independent rating agencies view auditor tenure as having a favorable impact on earnings quality. Finally, using earnings coefficients from a regression of one-year-ahead consensus earnings forecasts on reported earnings as a proxy for analysts' perceptions of earnings quality, we find evidence consistent with the hypothesis that financial analysts perceive earnings quality as improving with extended auditor-client relationships.

In general, most of our results are consistent with the hypothesis that audited financial statements, and in particular reported earnings, are perceived as more reliable for firms with longer auditor tenure. One implication of our results is that many capital market participants view longer tenure as having a favorable impact on audit quality. Overall, our study suggests that imposing mandatory limits on the duration of the auditor-client relationship might impose unintended costs on capital market participants. Our results are also subject to a caveat. Results from a regime without auditor term limits may not be applicable to a regime with mandatory auditor rotation because the incentives for auditors and clients might differ between the two regimes.

REFERENCES

Abarbanell, J., and V. Bernard. 2000. Is the U.S. market myopic? *Journal of Accounting Research* 38: 221–242.

Ali, A., and P. Zarowin. 1992. The role of earnings levels in annual earnings-returns studies. *Journal of Accounting Research* 30: 286–296.

American Institute of Certified Public Accountants (AICPA). 1978. *The Commission on Auditors Responsibilities: Report, Conclusions and Recommendations.* New York, NY: AICPA.

———. 1992. *Statement of Position Regarding Mandatory Rotation of Audit Firms of Publicly Held Companies.* SEC Practice Section. New York, NY: AICPA.

———. 1994. *Professional Standards.* New York, NY: Commerce Clearing House, Inc.

Ashbaugh, H., and M. Pincus. 2001. Domestic accounting standards, international accounting standards, and the predictability of earnings. *Journal of Accounting Research* 39: 417–434.

Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of audit experience. *The Accounting Review* 66: 218–239.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Steven. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73: 421–433.

———, D. Byard, C. Kile, and E. J. Reidl. 2002. High technology intangibles and analysts' forecasts. *Journal of Accounting Research* 40: 289–312.

Barth, M. E., and S. Kallapur. 1996. The effects of cross-sectional scale differences on regression results in empirical accounting research. *Contemporary Accounting Research* 13: 527–567.

———, J. A. Elliott, and M. A. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37: 387–413.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33: 173–190.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15: 1–24.

Bhojraj, S., and P. Sengupta. 2003. Effect of corporate governance on bond ratings and yields: The role of institutional investors and outside directors. *Journal of Business* 76: 455–475.

Bricker, R. 2002. Transparency, independence, technology and the CPA scope of services: New challenges to the profession. *The Ohio CPA Journal* (April–June): 48–50.

Brody, R. G., and S. A. Moscove. 1998. Mandatory auditor rotation. *National Public Accountant* (May): 32–36.

Brown, L., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-Year Earnings Forecasts and the Capital Market.* Studies in Accounting Research No. 21. Sarasota, FL: American Accounting Association.

———, R. Hagerman, P. Griffin, and M. Zmijewski. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9: 61–87.

Carmichael, D. R. 1999. In search of concepts of auditor independence. *The CPA Journal* (May): 39–43.

Collins, D. W., and S. P. Kothari. 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficient. *Journal of Accounting and Economics* 11: 143–181.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specialization. *Journal of Accounting and Economics* 20: 297–322.

Davis, L. R., B. Soo, and G. Trompeter. 2002. Auditor tenure, auditor independence and earnings management. Working paper, Boston College, Boston, MA.

Dechow, P. M. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18: 3–42.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17: 145–176.

———, and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25: 35–67.

———, and M. Hung. 2003. An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics* 35: 73–100.

Dhaliwal, D. S., and S. S. Reynolds. 1994. The effect of default risk of debt on the earnings response coefficient. *The Accounting Review* 69: 412–419.

Dopuch, N., R. R. King, and R. Schwartz. 2003. Independence in appearance and in fact: An empirical investigation. *Contemporary Accounting Research* 65: 83–113.

Duru, A., and D. M. Reeb. 2002. International diversification and analysts' forecast accuracy and bias. *The Accounting Review* 77: 415–433.

Easton, P. D., and T. S. Harris. 1991. Earnings as an explanatory variable for returns. *Journal of Accounting Research* 29: 19–36.

Elliott, R. K., and P. D. Jacobson. 1998. Audit independence. *The CPA Journal* (December): 30–37.

———. 2000. Testimony on Auditor Independence before the Securities and Exchange Commission. Available at: http://www.sec.gov/rules/proposed/s71300/testimony/elliott1.htm.

Fama, E. F., and K. French. 2000. Forecasting profitability and earnings. *Journal of Business* 73: 161–175.

Farmer, T., L. Rittenberg, and G. Trompeter. 1987. An investigation of the impact of economic and organizational factors in auditor independence. *Auditing: A Journal of Practice & Theory* 7: 1–14.

Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprise.* Statement of Financial Accounting Concepts No. 1. Norwalk, CT: FASB.

———. 1980. *Qualitative Characteristics of Accounting Information.* Statement of Financial Accounting Concepts No. 2. Stamford, CT: FASB.

Francis, J., and L. Soffer. 1997. The relative informativeness of analysts' recommendations and earnings forecast revisions. *Journal of Accounting Research* 34: 193–212.

Ghosh, A., Z. Gu, and P. C. Jain. 2005. Sustained earnings and revenue growth, earning quality, and earnings response coefficients. *Review of Accounting Studies* 10 (forthcoming).

Gramling, A. A., and D. N. Stone. 2001. Audit firm industry expertise: A review and synthesis of the archival literature. *Journal of Accounting Literature* 20: 1–29.

Greene, W. H. 1993. *Econometric Analysis.* Englewood Cliffs, NJ: Prentice Hall.

Gujarati, D. N. 1995. *Basic Econometrics.* New York, NY: McGraw-Hill.

Hogan, C. E., and D. C. Jeter. 1999. Industry specialization by auditors. *Auditing: A Journal of Practice & Theory* 18: 1–17.

Hope, O. 2003. Accounting policy disclosures and analysts' forecasts. *Contemporary Accounting Research* 20: 295–322.

Hunt, I. C. 2002. Testimony Concerning the Role of Credit Rating Agencies in the U.S. Securities Markets before the Senate Committee on Governmental Affairs. Available at: http://www.sec.gov/news/testimony/032002tsih.htm.

Johnson, V., I. K. Khurana, and J. K. Reynolds. 2002. Audit-firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19: 637–660.

Kaplan, R. S., and G. Urwitz. 1979. Statistical models of bond ratings: A methodological inquiry. *Journal of Business* 52: 231–261.

Kasznik, R., and M. F. McNichols. 2002. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40: 727–759.

Kinney, W. 1999. Auditor independence: A burdensome constraint or core value? *Accounting Horizons* 13: 69–75.

Kothari, S. P. 1992. Price-earnings regressions in the presence of pricing leading earnings: Earnings level versus changes specification and alternative deflators. *Journal of Accounting and Economics* 15: 173–302.

———, and R. Sloan. 1992. Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics* 15: 143–171.

Lang, M., and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71: 467–492.

Libby, R., and D. M. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* 22: 348–367.

Lundholm, R., and L. A. Myers. 2002. Bringing the future forward: The effect of disclosure on the returns-earnings relation. *Journal of Accounting Research* 40: 809–839.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security price changes. *Journal of Accounting and Economics* 13: 341–364.

Mansi, S. A., W. F. Maxwell, and D. P. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42: 755–793.

Mautz, R. K., and H. A. Sharaf. 1961. *The Philosophy of Auditing.* Monograph No. 6. Sarasota, FL: American Accounting Association.

Myers, J. N., L. A. Myers, and T. C. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78: 779–800.

Myers, L. A., and D. J. Skinner. 2002. Earnings momentum and earnings management. Working paper, University of Illinois at Urbana–Champaign and University of Michigan.

O'Brien, P. 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research* 28: 286–304.

PricewaterhouseCoopers. 2002. *Mandatory Rotation of Audit Firms: Will It Improve Audit Quality?* New York, NY: PricewaterhouseCoopers LLP.

Ryan, S. G., R. H. Herz, T. E. Iannaconi, L. A. Maines, K. Palepu, C. M. Schrand, D. J. Skinner, and L. Vincent. 2001. Commentary. SEC auditor independence requirements: AAA Financial Accounting Standards Committee. *Accounting Horizons* 15: 373–386.

Schipper, K. 1991. Analysts' forecasts. *Accounting Horizons* 4: 105–121.

———, and L. Vincent. 2003. Earnings quality. *Accounting Horizons* (Supplement) 17: 97–110.

Securities and Exchange Commission (SEC). 1988. *Financial Report.* Release No. 31. Chicago, IL: Commerce Clearing House.

———. 1994. *Staff Report on Auditor Independence.* Chicago, IL: Commerce Clearing House.

———. 2000. *Hearing on Auditor Independence.* July. Washington, D.C.: Government Printing Office.

———. 2001. *Final Rule: Revision of the Commission's Auditor Independence Requirements.* File No. S7-13-00. Washington, D.C.: Government Printing Office.

Shockley, R. A. 1981. Perceptions of auditor independence: An empirical analysis. *The Accounting Review* 56: 785–800.

Skinner, D. J., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7: 289–312.

Solomon, I., M. D. Shields, and R. O. Whittington. 1999. What do industry specialist auditors know? *Journal of Accounting Research* 37: 191–208.

Standard & Poor's. 1982. *Credit Overview: Corporate and International Ratings.* New York, NY: Standard & Poor's Corporation.

Stickel, S. 1989. The timing of and incentives for annual earnings forecasts near interim financial report releases. *Journal of Accounting and Economics* 11: 279–292.

Subramanyam, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22: 250–265.

———, and J. J. Wild. 1996. Going-concern status, earnings persistence, and informativeness of earnings. *Contemporary Accounting Research* 13: 251–273.

Teoh, S. H., and T. J. Wong. 1993. Auditor size and the earnings response coefficient. *The Accounting Review* 68: 346–366.

Turner, L. 2002. Testimony and Statements before the Committee on Banking, Housing, and Urban Affairs United States Senate, Available at: http://banking.senate.gov/02_02hrg/022602/turner.htm.

U.S. General Accounting Office (GAO). 2003. *Required Study on the Potential Effects of Mandatory Audit Firm Rotation.* Report to the Senate Committee on Banking, Housing, and Urban Affairs and the House Committee on Financial Services. Available at: http://www.gao.gov/new.items/d04216.pdf.

U.S. Senate. 1977. *Improving the Accountability of Publicly Owned Corporations and Their Auditors.* Prepared by the Subcommittee on Reports, Accounting, and Management of the Committee on Governmental Affairs. Washington, D.C.: Government Printing Office.

Van Horne, J. C. 1992. *Financial Management and Policy.* Englewood Cliffs, NJ: Prentice Hall.

Wall Street Journal. 2002a. Questioning the books: SEC queries of global crossing likely to resurface. (February 11): A8.

———. 2002b. Questioning the books: Court documents show Andersen ties with Enron were growing in early 1990s. (February 26): A6.

Warfield, T., J. Wild, and K. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20: 61–91.

Ziebart, D. A., and S. A. Reiter. 1992. Bond ratings, bond yields and financial information. *Contemporary Accounting Research* 9: 252–282.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 613–648

The Value Relevance of Financial Statements in the Venture Capital Market

John R. M. Hand
The University of North Carolina at Chapel Hill

ABSTRACT: This study examines the value relevance of financial statement data and nonfinancial statement information within and across the pre-IPO venture capital and post-IPO public equity markets. For a sample of U.S. biotechnology firms, I find that financial statements are highly value-relevant in the venture capital market, and that the signs of the associations between equity values and financial statement data in that market are similar to those in the public equity market, despite significant structural differences between the two. I also find that the value relevance of financial statements generally increases as firms mature, consistent with financial statements capturing the increasing intensity of assets-in-place relative to future investment options. In contrast, the value relevance of nonfinancial statement information decreases as firms mature, indicating that, in a dynamic sense, financial statements and nonfinancial statement information of venture-backed pre-IPO biotech companies are information substitutes in valuation, not complements.

Keywords: *biotechnology; financial statements; firm maturity; investment opportunities; nonfinancial statement information; private equity; value relevance; venture capital.*

Data Availability: *Key data items used in this paper were obtained from a proprietary source.*

I. INTRODUCTION

Much research has shown that financial statements are value-relevant in public equity markets (Barth et al. 2001; Kothari 2001). In this study, I break new ground by exploring the value relevance of financial statements in private equity markets—specifically, in the venture capital market. In the process, I shed light on the dynamic behavior of young firms' investment opportunity sets and on the dynamic relations of (and

---

The comments of two anonymous referees, R. T. Ball, M. Barth, R. Beatty, B. Beaver, R. Bushman, A. Davila, P. Dechow, G. Foster, W. Landsman, J. Raedy, D. Shackelford, K. R. Subramanyam, B. Trueman, H. Zhang, X-J. Zhang, and seminar participants at the University of California, Berkeley, Stanford University, The University of North Carolina at Chapel Hill, and University of Southern California are appreciated. I am grateful to M. Mumma and G. Kong of InterSouth Partners and F. Worthy of A. M. Pappas & Associates for sharing their insights into venture capital valuation, M. Edwards of Recombinant Capital for his valuation data, T. Seppä for sharing his time-series on the total amount of funding committed to U.S. venture capital funds, and S. Woodward of Sand Hill Econometrics for freely providing her proprietary private equity biotech valuation index. The Edward O'Herron, Jr. Fund for Distinguished University of North Carolina Faculty, the University of North Carolina Center for Technology and Entrepreneurial Venturing, and the University of North Carolina Center for Technology & Advanced Commerce provided generous financial support for this research.

Editor's note: This paper was accepted by Patricia Dechow, Editor.

Submitted June 2003
Accepted July 2004

between) the value relevance of financial statements and nonfinancial statement information as firms mature from their founding through three years after they go public.

The U.S. venture capital market is the part of the aggregate private equity market that consists of unregistered equity investments in young, very risky, and technology-oriented private firms by venture capital partnerships, corporations, pension funds, and university endowments (Gompers and Lerner 2000a). Although GAAP is the same for venture-backed private firms and publicly traded for-profit companies, the venture capital and public equity markets are structured very differently. The thesis of this paper is that despite these differences, common features of underlying firm economics and common GAAP combine to create predictable relations between equity values, financial statement data, and nonfinancial statement information within and across the pre-IPO venture capital and post-IPO public equity markets.

My first hypothesis is that since U.S. GAAP is designed to apply equally to all for-profit businesses, regardless of the market environment within which investors set prices, financial statements will, on average, be value-relevant in the venture capital market.

Second, although venture-backed technology-based firms are smaller, riskier, and more intangible-intensive than are the vast majority of firms in public equity markets, they engage in fundamentally the same kinds of financing, investing, and operating activities. I therefore hypothesize that the direction of the associations between financial statement data and equity values will be the same across the two markets—that is, the partial correlations between equity values and financial statement data will be of the same signs across the two markets.

Third, I hypothesize that financial statements will be value-irrelevant at the beginning of a firm's life but will become progressively more value-relevant as the firm matures. When a company of the kind that is funded by venture capital is founded, almost all its value is attributable to profitable expected future investment opportunities rather than assets-in-place. As a result, because GAAP does not allow the value of expected future investment options to be recorded in financial statements (only the relevant and reliable consequences of past transactions and activities may be included), a firm's financial statements will be value-irrelevant when it is founded. However, as it matures, the firm obtains financing, makes investments and undertakes operating activities, all of which convert potential investment opportunities into actual assets-in-place. Assuming that a maturing firm converts its existing investment options into assets-in-place faster than it discovers new ones, then its financial statements will reflect greater and greater fractions of its total equity value. Hence, as a firm matures its financial statements will become increasingly associated with its equity value.

Finally, I hypothesize that the value relevance of nonfinancial statement information will decrease with firm maturity. Nonfinancial statement information is defined as firm-specific information that is correlated with future investment opportunities (such as the number of patents issued to the firm) and non-firm-specific information that is correlated with the profitability of existing investments and/or industry- and economy-wide future investment opportunities (such as the level of equity values in the firm's industry). As firms mature, not only will the relative importance of assets-in-place and financial statements to their equity values increase, but also the relative importance of future investment opportunities and nonfinancial statement information to their equity values will decrease. Equivalently stated, in a dynamic sense financial statements and nonfinancial statement information are predicted to be information substitutes in valuation, not complements.

The data used to test these hypotheses are a specially constructed set of detailed data for an unbalanced panel of 204 U.S. biotechnology companies over the period 1992–2003.

Biotech firms are chosen because they are typically financed in well-defined stages by venture capital and go public quite quickly. The latter characteristic means that financial statement data spanning the entire life of the firm are often available in the IPO filing documents. My database exploits this by matching the pre-money venture capital valuations of a large set of biotech firms with up to five years of annual financial statement data in the IPO filings of those firms that filed to go public. This yields 481 pairs of venture capital firm valuations that are accompanied by the same firm's prior fiscal year's financial statements, spanning venture capital valuation points ranging from Series A to Series H financing rounds. This dataset is extended into the public equity market by incorporating data as of three months after each of the first three years following the IPO (for the subset of IPO filers that went public). Post-IPO stock price data are matched with the previous fiscal year's financial statements, yielding 449 public market observations.

Following prior research in entrepreneurial finance (Lerner 1994a; Gompers and Lerner 1999, 2000a, 2000b; Seppä 2003), hedonic log-linear regression models are used to assess the nature and strength of the associations between equity values, financial statements, and nonfinancial statement information. Two types of regressions are estimated—separate pooled regressions for each market, and financing-round-by-financing-round regressions within each market. The value relevance of financial statements is defined as the adjusted R^2 that can be uniquely attributed to financial statements in a regression of equity value on financial statement and nonfinancial statement information.

The financial statement variables included in the regressions are the major components of a firm's balance sheet and income statement, specifically, the firm's cash balance, noncash assets, long-term debt, annual revenues, cost of sales, SG&A, R&D expense, and unrecognized stock option dilution. Major components of balance sheets and income statements are used instead of aggregate book equity and net income in order to avoid the severe inferential distortions that can arise when evaluating the value relevance of financial statements of fast growing, highly intangible-intensive companies (Zhang 2001; Hand 2004).

The nonfinancial statement information included in the value-relevance regressions consists of variables that prior research in entrepreneurial finance has found to be correlated with the pre-money valuations of privately funded biotech companies, plus two previously unexamined factors. The former group consists of the AMEX Biotechnology Index (Lerner 1994a; Gompers and Lerner 2000b; Seppä 2003), the number and scope of patents filed by the firm (Lerner 1994a), the total number of the firm's upstream and downstream strategic alliances (Nicholson et al. 2003), an indicator for when the financing round was led by a corporate investor (Gompers and Lerner 1999), and indicators that measure the business stage the firm is at (Gompers and Lerner 2000b; Seppä 2003). The new factors that I examine are the Sand Hill Biotechnology Index of the value of privately held and privately financed biotech stocks, and the degree of equity dilution created by the current financing round.

Regression results are consistent with all four of the paper's hypotheses. First, biotech firms' financial statements are, on average, value-relevant in the pre-IPO, venture capital market. In pooled regressions, financial statements uniquely account for a highly significant 13 percent of the total regression adjusted R^2 of 76 percent. Nonfinancial information uniquely explains 18 percent, and 46 percent is variation common to both financial statement and nonfinancial statement data.

Second, the nature of the associations between equity values and financial statements in the venture capital market are similar to those observed in the public equity market. In both markets, with or without controlling for nonfinancial statement information, firms' equity values are positively related to cash, noncash assets, and R&D expense; unrelated to cost of sales; and negatively related to long-term debt and stock-option-related dilution. Of the coefficients on the eight financial statement variables included in the pooled regressions, only those on cash and revenues are reliably different across markets, both being larger in public equity markets.

Third, biotech firms' financial statements are value-irrelevant at Series A financings—the earliest formal venture capital financing round—but become progressively more value-relevant as firms mature. Consistent with the proposition that this is due to firms converting their future investment options into assets-in-place faster than they discover new ones, the ratio of total assets to firm value reliably increases as firms mature. These patterns are not a spurious effect of scale or other factors that can impede across-sample comparisons of adjusted R^2 (Gu 2002).

Fourth, financial statement data and nonfinancial statement information are informational substitutes rather than complements. As biotech firms mature and the value relevance of their financial statements increases, the value relevance of nonfinancial statement information decreases. The time-series correlation between the value relevance of financial statement data and the value relevance of nonfinancial statement information is reliably negative.

Last, I find that the value relevance of financial statement data and nonfinancial statement information are quite different at the IPO than expected from the maturity-based trends noted above. At the IPO, the value relevance of financial statement data dips temporarily below the maturity trendline, while the value relevance of nonfinancial statement information rises temporarily. Consistent with arguments made by finance research into other IPO pricing anomalies, I attribute this to the notion that IPOs are sold by sophisticated underwriters to unsophisticated retail investors who underweight the value relevance of financial statement data and overweight the value relevance of nonfinancial statement information.

In total, this paper contributes to several literatures. It expands the value-relevance literature by finding that financial statements in the pre-IPO venture capital subsector of the private equity market are qualitatively and quantitatively value-relevant in economically reasonable ways. This indicates that the concepts underlying U.S. GAAP are broad and robust. As intended by the FASB, GAAP "provide[s] information that is useful to present and potential investors ... in making rational investment, credit, and similar decisions" (FASB 1978, Concepts Statement No. 1, para. 34), regardless of whether the market in which investment decisions are made is private or public. This robustness exists despite large differences in price-setting mechanisms, the importance of assets-in-place versus future investment opportunities, and the level of regulation, liquidity, uncertainty, and information asymmetries between venture and public equity markets.

The paper also contributes to the accounting literature on the investment opportunity set. Prior work has found that in the cross-section, firms' investment opportunities are important determinants of many corporate decisions (Smith and Watts 1992; Gaver and Gaver 1993; Skinner 1993). This study demonstrates that not only do firms' investment opportunity sets decline as they mature, but also the decline is predictably related to both increasing value relevance for financial statements and decreasing value relevance for nonfinancial statement information. The increasing value relevance of financial statements as young firms mature also bookends the decreasing value relevance of financial statement data found by Barth et al. (1998) for firms approaching the ends of their lives in the form of sharply declining financial health and/or bankruptcy.

With respect to the accounting literature on the role of nonfinancial information in valuation, this paper contributes new knowledge beyond the dynamic relations observed between firms' investment opportunities and the value relevance of nonfinancial statement information. Specifically, although I concur with prior studies such as Amir and Lev (1996), Ittner and Larcker (1998), Deng et al. (1999), and Trueman et al. (2000) in their conclusions that nonfinancial information can be value-relevant, I find that for biotech firms, nonfinancial statement information is only materially value-relevant prior to filing for an IPO. In the post-IPO market, nonfinancial statement information seems to be almost entirely value-irrelevant once the value relevance of financial statements is controlled for. Moreover, I find that financial statement data and nonfinancial statement information are substitutes, not complements, in valuation. This runs counter to the findings of Amir and Lev (1996).

Finally, this paper contributes to entrepreneurial finance, which has not previously examined the relevance of financial statements. I show that financial statements are highly related to pre-IPO firm equity values. I also show that when venture capital valuations are pooled over all pre-IPO financing rounds, nonfinancial statement information of the type examined previously by venture capital researchers still has explanatory power once financial statement data are controlled for. I further identify two new value-relevant items of nonfinancial statement information—the level of biotech company valuations in the entire venture capital market, and the degree of new equity dilution created by the current round of funding. The results regarding dilution strongly suggest that the demand for equity is far from perfectly elastic in the venture capital market. In addition, not controlling for financial statements leads to the erroneous conclusion that the value relevance of nonfinancial information does not change as the firm matures, when, in fact, it declines sharply.

The remainder of the paper proceeds as follows. Section II describes the institutional background of the U.S. private equity market, particularly venture capital, and provides a sketch of biotechnology and business research into biotech firms. Section III develops the value-relevance hypotheses. Section IV explains the sample selection, describes the database employed, and reports descriptive statistics. Section V motivates the regression models that are used. Section VI reports regression results along with the findings from various robustness diagnostics and notes the key limitations of the study. Section VII concludes.

II. THE VENTURE CAPITAL MARKET AND BIOTECHNOLOGY

Private Equity and Venture Capital

There are four types of private equity: organized, angel, informal, and Rule 144A (Fenn et al. 1995).[1] Venture capital is the subpart of organized private equity that invests in young, typically high-technology, firms. Venture capital funds are usually structured as partnerships that raise money in staged amounts from companies, institutions, and wealthy private investors. A fund generally has a ten-year life and invests in a portfolio of private companies, often focused in one or two sectors such as biotechnology or software (Sahlman 1990). Although venture capitalists put up only about 1 percent of the limited partnership's capital, they manage the fund through their role as the general partners. In exchange for finding, screening, and deciding upon which companies to invest in, they are paid an annual management fee that is usually between 1.5 percent and 3 percent of the fund's committed capital or net asset value, and receive about 20 percent of the profits made by the fund's investments. The typical investment made by a venture fund in a company is in the form of illiquid preferred stock that only converts into liquid common stock or cash at one of two major exit points—an IPO, or the sale of the company to another entity. Although venture funds provide a firm with a measure of long-term financing through their investments in several financing rounds, they also supply business expertise and connections. A partner from the venture fund usually serves on the firm's board of directors; provides financial sophistication, operating services, and business contacts; helps recruit key personnel; and imparts financial and strategic discipline.

---

[1] Organized private equity is unregistered investments in the equity of private companies, made either directly by professional investors such as corporations, pension funds, and endowments or indirectly by these investors through intermediaries, particularly venture capital partnerships. Angel private equity consists of investments by wealthy individuals, typically arranged by matchmakers such as lawyers and accountants (Wong 2002). Informal private equity is similar to angel capital except that firms sell unregistered equity securities to both institutional investors and wealthy individuals across a larger number of such investors. Rule 144A private equity is private equity offerings underwritten under the SEC's Rule 144A, which establishes the rules and conditions under which private securities are permitted to be traded among certain classes of institutional investors (Fenn et al. 1995).

Distinctives of the Venture Capital Market

The venture capital market differs from the public equity market in its institutional structure and the type of firms that it finances (Wright and Robbie 1998; Gompers and Lerner 2000a). In terms of the number, valuation, and diversity of firms involved, the venture capital market is much smaller than is the public equity market. Whereas the public equity market is highly regulated by the SEC and stock exchanges, the venture capital market is not. Firm valuations in the venture capital market are set through face-to-face negotiations between management and a small number of wealthy, professional, and risk-tolerant investors. Valuations in the public equity market are set anonymously without direct contact with management via the interactions of large numbers of investors, many of whom are risk-averse and do not have significant wealth or professional investing experience. Unlike in public markets, the valuations of individual venture-backed firms are difficult and costly for nonparticipating investors to observe. Trading is cheap and all but continuous in the public equity market, but it is expensive and rare in the venture capital market. The trading that does occur in the venture capital market is almost entirely in the form of long investments in new equity, not secondary transactions in existing shares. Venture capital investments are also far less divisible than are public equities. Short selling is feasible and frequent in the public market, but nonexistent in the venture capital market. Venture capital investors are more likely to be able to extract management's private information because of the board positions they hold and their frequent interactions with and monitoring of management. In contrast, public equity investors typically engage in indirect monitoring, rely primarily on public data, and face significant information asymmetries. In total, the venture capital market is illiquid, incomplete, and imperfect—to a much greater degree than is the public equity market.

Research into Venture-Backed Companies

Research into venture capital has blossomed over the past decade. An excellent summary is provided by Gompers and Lerner (2000a), spanning topics such as the compensation of venture capitalists (Gompers and Lerner 1999); the optimal investment, monitoring, and staging of venture capital (Gompers 1995); the decision to go public (Lerner 1994b); and the long-run performance of venture-backed IPOs (Brav and Gompers 1997). Most recent work includes that of Kaplan and Stromberg (2003, 2004), who empirically relate the characteristics of venture capital contracts to theories of financial contracting, and Seppä (2003), who examines the relations between investment syndication and the efficiency of venture capital firms, as well as the role of network positions in determining the performance of venture capital firms.

Biotechnology and Biotechnology Firms

Biotechnology is the application of technology to the life sciences, wherein living cells or their processes are used to solve problems and perform specific industrial or manufacturing processes. Biotech applications include the production of drugs, synthetic hormones, and bulk foodstuffs; the bioconversion of organic waste; and the use of genetically altered bacteria. Biotech firms are highly dependent on the intellectual property (ideas, discoveries, patents) generated through their large R&D expenditures and, as such, are among the most intangible-intensive of businesses. The value chain of the typical biotech firm stretches some 10–15 years from founding through patenting to successful FDA approval and product sales (see Figure 1). Biotechnology is therefore a very risky but potentially very lucrative equity investment.

The prototypical young biotech firm is in an intense R&D race against competitors to discover and patent a new drug. As such, it has large capital needs over a long period of time. In the early stages of life, the firm's capital needs are met by venture capital and strategic equity investments from pharmaceutical companies. However, capital needs typically become so large that they can only be satisfied through an IPO or a buyout by a large pharmaceutical company. Successful biotech firms therefore tend to go public rapidly, and it is not uncommon for a biotech firm's S-1 filing with the SEC to contain financial statements that span its entire life.

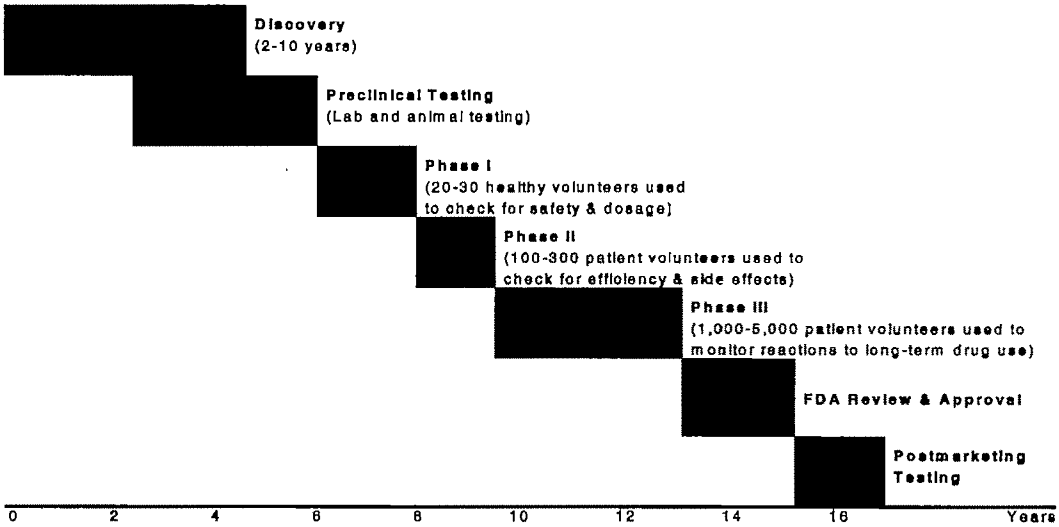
Economic research into biotechnology has spanned three areas: intellectual capital (Zucker and Darby 1996; Zucker et al. 1998), strategic alliances (Robinson and Stuart 2002), and valuation (Stuart et al. 1999; Nicholson et al. 2003; Darby et al. 1999). In accounting, Joos (2002) finds that the level and rate of growth in R&D expense, R&D success, and competitive structure explain cross-sectional variation in market-to-book ratios for pharmaceutical drug manufacturers. Ely et al. (2003) conclude that the average FDA stage of a firm's portfolio of drugs conditions the value relevance of the firm's R&D expenditures. Guo et al. (2004) find that biotech firms' disclosures affect their bid-ask spreads and stock return volatility. Hand (2004) reports that the mapping of publicly traded biotech firms' R&D expenditures into equity market value is a function of the location of R&D in the biotech value chain of discovery, development, and commercialization, as well as the growth rate in R&D spending.

III. HYPOTHESES

The goals of this paper are to develop and test hypotheses about the value relevance of financial statement data and nonfinancial statement information within and across the venture capital and public equity markets. The thesis of the paper is that despite significant economic and structural differences between the two markets, common features of underlying firm economics and common GAAP combine to create predictable relations between equity values, financial statement data, and nonfinancial statement information.

The first hypothesis stems from the fact that the FASB's Conceptual Statements and Issued Standards are designed to be useful to investors in all types of for-profit businesses, not just publicly traded entities. Although firms in the venture capital market have more concentrated ownership and are smaller, riskier, and more intangible-intensive than firms in the public market, and although it has been argued that U.S. GAAP was developed for tangible-intensive businesses and is therefore less well suited to intangible-intensive technology firms (Lev and Zarowin 1999; Lev 2001), it seems unlikely that financial statements would be entirely without value relevance for all companies at all points prior to filing for an IPO. This is because all firms in both markets engage in financing, investing, and operating activities—only on different scales and to different degrees. The first hypothesis is

FIGURE 1
Biotechnology Value Chain



Source: Ernst & Young LLP, *Biotechnology Industry Report Convergence*, 2000.

Discovery and Preclinical Testing: The drug development process usually begins with the scientific discovery of a gene or other biological pathway involved in a disease. Discovery can take 2–10 years. From discovery, a target for therapeutic intervention is established. Preclinical tests are conducted in the lab using individual cells or sometimes animals to evaluate the safety and potential for effectiveness in humans. If the target is determined to be legitimate, the company files an Investigative New Drug (IND) application with the FDA for clearance to begin testing on humans. Even after these first few years of research and testing, however, most new drug candidates will never make it to the market.

Phase I Trials: Human testing begins. The purpose of a Phase I trial is to use a small number of volunteer patients to establish basic safety and maximum dosage parameters.

Phase II Trials: This stage of clinical study is much more involved, requiring many months to plan, set up, and recruit trial participants. Phase II is conducted on a larger group of patients with the targeted disease to study the efficacy of the drug at various doses and confirm its safety. They typically use blinding and placebo controls to achieve scientifically sound results. Phase II often lasts two years, and sometimes a drug will undergo multiple Phase II trials for different indications (for example, to treat different types of cancer). This may be the most critical phase in terms of sorting winners from losers. As a rule of thumb, drugs that complete Phase II and move on to Phase III have about a 50 percent success rate of reaching the market, though some studies suggest the rate is higher.

Phase III Trials: These tests are designed with a specific endpoint—a measurable result that clearly demonstrates success in combating the targeted disease. The endpoint must be agreed upon by the FDA as an outcome that will lead to marketing approval. The trial involves a large group from the targeted patient population and uses controls such as double-blinding (neither patient nor doctor knows who is getting a placebo). Multicenter trials are common to show that results are reproducible when administered in different clinical settings. This pivotal phase often lasts two to three years from initial design to study completion, and here again it is common for drugs to undergo more than one Phase III trial for different indications or to support different therapy combinations.

FDA Approval Process: If a drug successfully completes Phase III, then the company gathers all of its clinical data and files an application for marketing approval with the FDA. It often takes three to six months just to prepare the application. Another six to 12 months can pass before an FDA advisory panel reviews the application and makes a recommendation. This advisory panel has expertise in the drug's specific area of therapeutic or disease characteristics, and its recommendation for denial or approval is normally followed by the FDA (though another six to 12 months can pass before that happens).

Modified from an article by James Hale, available at:
http://www.theonlineinvestor.com/industries.phtml?content=is_bio2.

therefore that the financial statements will on average be value-relevant in the venture capital market.

The second hypothesis extends this line of reasoning. Because U.S. GAAP does not depend on whether a for-profit firm is privately held or publicly traded, I hypothesize that the nature of the associations between equity values and financial statement data in the venture capital market will accord with economic intuition and will be the same as those in the public market. That is, I predict that the partial correlations between equity values and key financial statement items will have economically sensible signs, and that the signs will be the same across markets. For example, the GAAP definition of noncash assets leads to the prediction that the partial correlation between firms' equity values and noncash assets will be positive, not negative, but that the converse will be true for a liability such as long-term debt.

The third hypothesis is that financial statements will be value-irrelevant at the beginning of a firm's life, but will become progressively more value-relevant as the firm matures. As highlighted by Miller and Modigliani (1961), the value of a firm's equity can be broken into two parts: the present value of net assets-in-place as a result of past investments, and the present value of future profitable investments or growth opportunities. Growth opportunities are the real options that a firm has or may create to make future investments that earn a rate of return in excess of its opportunity cost of capital (Myers 1977). Growth opportunities are frequently referred to as the firm's "investment opportunity set" and have been found to be an important determinant of many corporate decisions (Adam and Goyal 2003).[2]

When a firm is founded, its value consists almost exclusively of ideas the founders have for profitable future investments—*viz.*, investment/growth opportunities—rather than on-balance-sheet assets-in-place (Myers 1977). As the firm matures, its investment opportunities are financed and converted into tangible and intangible assets and liabilities recorded in financial statements through double-entry bookkeeping (e.g., cash, PP&E, R&D, trademarks, debt). One result of this process is that the fraction of the firm's equity value attributable to assets-in-place increases relative to that which is attributable to future investment opportunities.[3] This view is consistent with the conventional start-up/growth/maturity/decline model of the lifecycle of a firm (Smith et al. 1985; Black 1998).[4]

U.S. GAAP, which applies equally to private and public for-profit firms, does not allow the value of expected future investment opportunities to be recorded in financial statements, only the relevant and reliable assets and liabilities associated with past investments. This means that as a firm matures and the fraction of value attributable to its assets-in-place increases relative to that of its future investment opportunities, the firm's financial statements will reflect a greater and greater fraction of the firm's value. A key manifestation of this linkage will therefore be that the value relevance of a firm's financial statements will increase as the firm matures.

---

[2] For example, the relative mix of assets-in-place versus future investment opportunities affects a firm's accounting policies (Skinner 1993), capital structure (Myers 1977; Smith and Watts 1992), compensation contracts (Smith and Watts 1992; Gaver and Gaver 1993), and dividend policy (Smith and Watts 1992).

[3] Although it is possible that a maturing firm could maintain a nondecreasing ratio of value from assets-in-place relative to value from future investment opportunities, this would require the firm to continually create new and increasingly large future investment opportunities. In expectation, this is economically unlikely.

[4] In the start-up phase there are few assets-in-place and the largest fraction of a firm's value stems from its ideas, intellectual property, and growth opportunities. In the growth phase, financing has been obtained, investments have been made, and operating activities are underway. Although the fraction of firm value attributable to assets-in-place is higher than in the start-up phase, growth opportunities remain large and near-term financing requirements remain substantial. In the maturity phase of the life cycle, the firm's growth has slowed substantially, its assets-in-place dominate its valuation, and its cash from operations is strong enough that it does not need external financing. Finally, in the stagnant or decline phase, the firm is either in a no-real-growth steady state where its earlier abnormal profits have all been competed away, or the firm is destroying shareholder value in its investments and operations, making it a candidate for restructuring, a takeover, or liquidation (Smith et al. 1985).

My fourth hypothesis stems from the observation that financial statements are not the only source of value-relevant information. Specifically, I propose that as a firm matures, the value relevance of its nonfinancial statement information will decrease. Nonfinancial statement information is defined as firm-specific data that are correlated with future investment opportunities (e.g., the number of patents issued to a firm) as well as non-firm-specific information that is correlated with the profitability of existing investments and/or industry- and economy-wide future investment opportunities (e.g., the level of equity values in a firm's industry). As firms mature and convert investment opportunities into assets-in-place, not only will the relative importance to equity values of assets-in-place and financial statements increase, but also the relative importance of future investment opportunities and nonfinancial statement information will decrease. That is, I predict that in a dynamic sense financial statements and nonfinancial statement information will be informational substitutes in valuation, not complements. It should be noted that this fourth hypothesis does not automatically follow if the third hypothesis is true. It could logically be the case that the value relevance of both financial statement data and nonfinancial statement information both increase as firms mature.

IV. DATA

This section describes the construction and characteristics of a database that combines private equity valuations, financial statement data, and nonfinancial statement information. The strengths of the database are that it is large, not unduly clustered in time, and rich in different kinds of information. It contains a wide variety of longitudinal financing points (Series A through Series H financing rounds, as well as many nonfinancing public valuation points) made by firms that either successfully or unsuccessfully tried to go public. The database has two main limitations. It does not contain venture capital valuations for firms that did not file to go public because they went bankrupt, merged, or chose to remain private. If financial statement data are more value-relevant for companies that were successful in that they filed to go public, then there may be an upward bias in the measured pre-IPO value relevance of financial statements. This concern is addressed in Section VI. Also, all of the IPOs from which financial statement data are taken were filed between 9/29/95 and 12/14/01, which is only approximately a six-year period.

Equity Valuation Data

Venture capital firm valuations were purchased from Recombinant Capital (Recap, visit http://www.Recap.com), a West Coast consulting firm that specializes in collecting and selling information on biotechnology companies. Recap has gathered what it indicates is a full set of round-by-round financings for each of more than 600 biotech companies beginning in the early 1980s. Recap's equity valuation histories come mainly from IPO filing documents and SEC filings by companies that acquire biotech firms.[5] The valuation histories are based on primary preferred and common share issuances from the date of founding until the most recent equity financing or other share issuance (which may be after a successful IPO). The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. Preferred shares are converted into common at the ratios specified in the financing agreements.

---

[5] The data items in Recap's valuation file consist of company name; the type of financing round (e.g., Founding, Series A, IPO, etc.); the category of the round (e.g., Founder/Insider, Venture, Corp. Private, etc.); the date the round was filed (that is, begun); the date the round was concluded; the amount of dollars raised in the round, if any; the price per share of that round's financing; the post-money valuation; the number of shares involved (although this is infrequent); and the names of up to three investors, with the first investor being the lead investor in the round.

Post-IPO firm equity values were obtained from standard online sources. Three post-IPO valuation dates were used, namely three months after the first, second, and third fiscal year-ends following the IPO offering date (if an IPO occurred, since not all biotech firms in the sample that filed to go public successfully accomplished an offering).

Financial Statement Data

For the subset of firms in Recap's database that filed to go public, pre-IPO financial statement data were collected from S-1 and 424B documents when they were available online at http://www.sec.gov. This exploits the fact that when a firm files to go public, it has to provide five years' worth of audited (albeit coarse) historical financial statement data.[6] Post-IPO financial statements were taken from 10-Ks.

In the pre-IPO part of the total dataset, on a firm-by-firm basis, each year of pre-IPO financial statement data was (where possible) matched with the first—and only the first—venture capital pre-money valuation following the fiscal year-end, providing that the valuation date was less than a year after the fiscal year-end.[7] No set of annual financial statements was matched with more than one pre-money valuation. This process yielded 481 pairs of pre-money venture capital valuations for which there were financial data from the preceding fiscal year, covering a total of 193 firms. To be considered usable, a firm's cash balance, SG&A expense, and R&D expense had to be positive.[8] In the post-IPO part of the database, for each biotech firm that successfully went public, its equity market value was (where possible) noted three months after each of the first three fiscal year-ends after its IPO. These equity market values were then matched with the preceding fiscal year's financial statement data, yielding 441 pairs of public market valuations for which there were financial data from the preceding fiscal year, covering a total of 160 firms.

The financial statement items included in the database were three primary components of total shareholder equity—cash, noncash assets, and long-term debt—and the main components of biotech firms' net income, namely annual revenues, cost of sales, SG&A costs, and R&D expense.[9] Long-term debt includes capitalized lease obligations. Revenue is the sum of collaborative, contract, grant, license, research, and product revenues. In addition to these recognized numbers, the degree of stock option dilution was included, defined as the number of shares under option divided by the number of shares outstanding. Stock option dilution is a proxy for economic cost of stock options that is almost never recognized in firms' financial statements. Stock option data are sometimes missing because firms are only required to provide three years of pre-IPO filing stock option data in their S-1 or 424B-4 filings. To account for this, an indicator variable is included in the regressions and is set equal to 1 if stock option data are missing, and 0 otherwise. Missing stock option dilution observations are then reset to 0.

---

[6] I am unaware of any other systematic, large-sample source of financial statement data for pre-IPO companies. Venture capital funds contacted during the course of this research indicated that they keep imperfect historical records of the financial statements of companies they invest in. Noting that "accounting data for private firms is unavailable," Gompers (1995) uses SIC industry averages from Compustat for each venture capital-funded firm in his analysis of the determinants of the duration of, and funding amounts provided to, venture capital-backed firms.

[7] For example, if a firm with a calendar year-end filed to go public on 11/5/98, and had Series A, B, and C rounds on 1/6/96, 10/27/96 and 12/5/97, respectively, the Series A and Series C would be usable, but the Series B would not because it is the second financing in 1996. In my design, only the first financing in 1996 can be matched with financial statements for the year ended 12/31/95 because it not possible to know how the financing on 1/6/96 affected financial statements as of 10/27/96. The strength of this approach is that no private equity market valuations are matched to what I will call "financing contaminated" financial statements. The limitation is that the gap between the valuation financing date and the preceding fiscal year end can be as little as a few days and as long as almost a year. However, the noise created by this weakness is unlikely to create regression biases in favor of my hypotheses.

[8] A biotech firm that has no cash, or no SG&A, or no spending on R&D, is highly unusual. Approximately 2 percent of the initial observations were deleted as a result of these restrictions.

The GAAP definitions of assets and liabilities lead to the prediction that in a regression of equity values on financial statement data, the coefficients on cash and noncash assets in both venture capital and the public equity markets will be positive, while the coefficients on long-term debt and stock option dilution will be negative. For income statement variables, I expect to observe a positive coefficient on revenues to the extent that revenues are not entirely transitory, a positive coefficient on R&D because the bulk of the benefits from R&D emerge in future periods, and a negative coefficient on cost of sales.[10] I make no sign prediction on SG&A, because SG&A is a mixture of period expenses that would be expected to be negatively related to equity value, such as the rent on the firm's facilities, and certain costs that provide future benefits, such as salaries for senior management and key scientific personnel, which would be expected to be positively related to equity value due to their multiperiod nature.

The regressions that I estimate characterize firms' equity values as a function of the major components of their book equity and net income rather than their book equity and net income per se, because young biotech firms make highly intensive and rapidly growing investments in R&D and therefore have significantly biased financial statements. Zhang (2001) demonstrates analytically that accounting conservatism combined with rapid growth in intangible assets can dramatically distort the associations between aggregate financial statement data and equity value. If the firm's investments in intangible assets are sufficiently intense, reported losses will be negatively, not positively, associated with equity values. Hand (2003) confirms Zhang's (2001) predictions for intangible-intensive Internet companies and demonstrates that the distorted relations between equity values and aggregate financial statement data for intangible-intensive firms can be avoided by replacing book equity and net income with their key components, such as the major categories of assets, liabilities, revenues, and expenses. This prevents the associations between intangible assets/expenses and equity value from contaminating the associations between tangible assets/expenses and equity value.

---

[9] By focusing on major components, this approach omits minor components. On the balance sheet, total assets = cash + noncash assets = total liabilities + convertible preferred stock + stockholders' equity. Of total liabilities, I include only long-term debt (defined as including long-term notes payable) because current liabilities are not identified in the five-year historical summary of coarse financial statement data in S-1 filings. If the sum of current liabilities and noncurrent liabilities outside of long-term debt is estimated from the data that are available, and is included in the value-relevance regressions, no inferences are changed. In biotech firms' income statements, net income = revenues − cost of sales − SGA − R&D + interest income − interest expense − income tax expense +/− miscellaneous items. Interest income and interest expense are excluded from the value-relevance regressions because they are directly and causally related to firms' cash holdings and long-term debt. Income tax expense is omitted because it is rarely present, as are miscellaneous items. Net income attributable to common shareholders = net income − accretion to redemption value and/or accrued dividends related to beneficial conversion features of preferred stock. The latter variable is not included in the regressions because the corresponding preferred stock is effectively included in the selection of the balance sheet items cash, noncash assets, and long-term debt (since cash + noncash assets − long-term debt = current liabilities + long-term liabilities outside of long-term debt + preferred stock outside of stockholders' equity + stockholders' equity).

[10] Because cost of sales is only recorded for product sales, and young biotech firms typically earn revenues from several other sources before they are mature enough to generate product sales (collaborations, contracts, grants, licenses, and research), few firms in the venture capital market portion of the database reported cost of sales.

Nonfinancial Statement Information

A substantial amount of nonfinancial statement information was collected for this study in order to include in the value-relevance regressions many of the variables that prior work has found to be value-relevant for venture-backed biotech firms. Beyond these, I also explore the value relevance of two new nonfinancial statement factors. The nonfinancial statement information that was collected comprises the following items:

(1) The level of the AMEX Biotechnology Index on the valuation date. Lerner (1994a), Gompers and Lerner (2000b), and Seppä (2003) find that biotech firms' pre-money valuations are larger the higher are overall publicly traded biotech equity values, because overall publicly traded biotech equity values provide a strong indication of the economic prospects of the biotech sector as a whole.[11]

(2) The age of the firm. Gompers and Lerner (1999), Nicholson et al. (2003), and Seppä (2003) find that the older biotech firms have larger pre-money valuations. They attribute this to older firms being less risky.

(3) The number and scope of the patents filed by the firm on the valuation date. Lerner (1994a) argues that intellectual property is a young biotech company's most valuable asset. To test this proposition, he computes the number of patents issued to a firm at the valuation date and develops a proxy for patent scope using the International Patent Classification scheme.[12] He finds that both the number and scope of biotech firms' issued patents are reliably positively associated with pre-money valuations. Lerner (1994a) also develops a proxy for patent uniqueness, which for parsimony I do not include in this study. I adopt the same general approach as Lerner (1994a), except that I use patents filed, rather than patents issued, as of the valuation date. Results are not sensitive to this distinction, but conversations with venture capitalists and patent lawyers specializing in biotechnology indicate that the average lag of about three years between when a patent is first filed and when it is issued leads venture capitalists to significantly weight patent filings. Patent data were extracted from http://www.uspto.gov.

(4) The total number of the firm's upstream and downstream strategic alliances. Consistent with the view that alliances with pharmaceutical companies are credible endorsements of a biotech company's prospects, Nicholson et al. (2003) report that biotech companies that sign alliance deals with pharmaceutical companies (so-called "downstream" deals) are valued significantly higher by venture capitalists. I expand this type of analysis to include "upstream" deals, such as that between

---

[11] The Amex Biotechnology Index (BTK) is an equal-dollar-weighted index designed to measure the performance of a cross-section of companies in the biotechnology industry that are primarily involved in the use of biological processes to develop products or provide services. Such processes include, but are not limited to, recombinant DNA technology, molecular biology, genetic engineering, monoclonal antibody-based technology, lipid/liposome technology, and genomics. The BTK Index was established with a benchmark value of 200 on October 18, 1991. It is rebalanced quarterly based on closing prices on the third Friday in January, April, July, and October to ensure that each component stock continues to have approximately equal weight in the index.

[12] For an explanation of the International Patent Classification scheme, and how it differs from the U.S. classification scheme, see Lerner (1994a, 320–321).

a biotech firm and a university. I predict that the greater the number of a firm's total alliances, the higher will be its equity value. Alliance data were obtained from http://www.recap.com.

(5) An indicator for whether a private equity financing round was led by a corporate investor. Gompers and Lerner (1999) find that corporate investors assign higher pre-money valuations when investing in private equity financings than do venture capital funds, either because corporations perceive there to be strategic synergies available, or because corporations typically engage in less monitoring and business development than do venture capital funds and therefore earn a lower expected return on their investment. Data on whether a venture capital financing round was led by a corporate investor were obtained from http://www.recap.com.

(6) Indicator variables that measure the business stage of the firm. Gompers and Lerner (1999, 2000b) and Seppä (2003) control for firm maturity using firm age and either the round number (e.g., first, second, third, etc.) or the firm's business stage (start-up, development, beta, shipping, profitable, or restart). Unlike the VentureOne database used in most prior studies, the Recap database codes the Series of the financing round—whether the round is Series A, Series B, etc.—rather than the firm's business stage. Although I do not make predictions on the signs of Series indicators, they play a important role in mitigating the possible selection bias arising from the fact that only firms that file for an IPO are included in the sample. This is because selection bias is likely to be more severe across rounds than within rounds. Financing round indicators also help control for discount rates that are expected to decline as firms mature.

Beyond these, I also examine the value relevance of two new nonfinancial statement factors:

(7) The Sand Hill Biotechnology (SHB) Index of the value of privately held and financed biotech stocks. The SHB Index is a proxy for the biotech industry's overall prospects, separate from that of the AMEX Biotech Index of publicly traded biotech stocks. I therefore expect it to be positively related to equity values. As described by its developer, Sand Hill Econometrics of Palo Alto, the SHB Index is a proprietary, unbiased index of the venture capital value of pre-IPO privately funded biotech companies. The SHB Index had a benchmark value of 100 in January 1989. Using bias-corrected data on company values, Sand Hill Econometrics builds the Index with tools akin to those used for repeat-sales indices of houses and other assets.[13]

(8) The degree of equity dilution created by the current financing round, defined as the number of shares issued to new investors divided by the number of pre-money shares outstanding. This variable measures the degree to which the venture capital market is imperfect in that the demand for equity is less than perfectly elastic. A negative coefficient on dilution implies that demand for the firm's equity is less

---

[13] According to Sand Hill Econometrics, in dollar terms the Index covers virtually all private biotech companies in the United States that have raised money from outside investors since 1987. The valuation events used to construct the Index include funding rounds, IPOs, acquisitions, liquidations, and shutdowns. The Index addresses the measurement problems unique to venture investing, namely intermittent pricing and selection bias. The Index is built from individual company data, not venture-fund data, because company-level data allow for the measurement of value at a point in time and for statistical correction for selection bias. Sand Hill Econometrics' methods correct for bias in the reporting of value for known deals and for bias in the likelihood of companies falling out of available data sources. Companies doing poorly are less likely to report value and more likely to be left out of subsequent systematic capture of funding round or shutdown information.

than perfectly elastic because the price at which new investors are willing to invest is decreasing in the number of new shares to be issued. Venture capital markets may be imperfect in this way because the availability of strong substitutes for any given equity investment are more limited than in public equity markets, given the much smaller size of the venture capital market and the very illiquid nature of venture capital investments. Prior work by Gompers and Lerner (2000b) and Seppä (2003) has also examined whether venture capital valuations were affected by demand, using the total amount of funding committed to U.S. venture capital funds in the 12 months leading up to the valuation date. Both studies found evidence that inflows of capital into venture funds increase the valuations of funds' new investments; that is, firm valuations get inflated because "too much money is chasing too few deals." The advantage of using dilution created by the current financing round instead of annual inflows of capital into venture funds as a proxy when testing for less than perfectly elastic demand is that dilution created by the current financing round is firm-specific. If the demand for equity in the venture capital market is indeed less than perfectly elastic, then the dilution created by the current financing round will be negatively associated with pre-money valuations. Moreover, if dilution created by the current financing round is a superior proxy, it should also statistically dominate the association between equity value and annual inflows of capital into venture funds.

Descriptive Statistics

Table 1 provides descriptive statistics on the major financial statement and nonfinancial statement information items for the sample biotech firms. Private equity financings are evenly spread over the ten-year period 1992–2001, and public market valuation dates over the eight-year period 1996–2003 (Panel A). The distribution of valuation observations across the venture capital and public equity markets is such that the vast majority of sample firms are represented in both the venture capital market and the public equity market datasets, rather than only one of the two (Panel B). Of the total sample of 204 firms, 193 have at least one valid valuation in the venture capital market, and 160 have at least one valid valuation in the private equity market. Also, 85 percent (136/160) of firms in the private equity market dataset have valuations for each of the maximum of three post-IPO years, while 82 percent (158/193) of firms in the venture capital market dataset have two or more valuations.

Most biotech firms file to go public quite rapidly, mainly between four and six years after they were founded (Panel C). The IPO filing date for the median firm occurs during preclinical testing and/or Phase I trials (see Figure 1), although a few firms take much longer to file to go public. The majority of sample companies are in pharmaceutical preparations, SIC 2834, or commercial, physical, and biological research, SIC 8731 (Panel D). Despite this concentration, sample firms comprise a total of 22 different four-digit SIC codes. Of firms, 47 percent had their headquarters at the time of the IPO filing in California and 14 percent were in Massachusetts (Panel E). Headquarters clustering arises because California and Massachusetts contain many top universities with star scientists on their faculties (Zucker et al. 1998).

Table 2 reports the median values of various financial statement and nonfinancial variables in private and public markets on a round-by-round basis. Table 2 is organized from left to right in the typical sequence followed by a firm funded by private equity, beginning with the Series A financing. The number of observations at a given private equity round differs from the number of firms in the sample because not all firms file to go public after their Series G or H round; some go public as early as after their Series A round. In addition, a given set of annual financial statements is only ever matched to one financing round. For example, if a calendar-fiscal-year firm had a Series B financing on 1/10/96 and a Series C financing on 12/28/96, only the Series B valuation—matched to the 12/31/95 financial

TABLE 1

Descriptive Statistics on 481 Venture Capital Market Valuations and 449 Public Equity Market Valuations for an Unbalanced Panel of 204 U.S. Biotechnology Firms, 1992–2003

Panel A: Distribution of Valuation Dates

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Venture	15	29	46	60	77	70	67	60	52	5	0	0
Public	0	0	0	0	8	49	71	71	40	78	74	58

Panel B: Distribution of Valuations per Firm across the Venture Capital and the Public Equity Markets

		# Valuations per Firm in the Public Equity Market					
		0	1	2	3	Sum	Weighted Average
# Valuations Per Firm in the Venture Capital Market	0	0	0	1	10	11	2.9
	1	6	1	3	25	35	2.3
	2	18	1	4	38	61	2.0
	3	16	4	6	39	65	2.0
	4	4	0	4	22	30	2.5
	5	0	0	0	2	2	3.0
	Sum	44	6	18	136	204	
	Weighted Average	2.4	2.5	2.5	2.3		

Panel C: Number of Years between Firm Founding and IPO Filing

Years	# Firms	Years	# Firms
0 to 1 year	0	7 to 8 years	14
1 to 2 years	5	8 to 9 years	15
2 to 3 years	12	9 to 10 years	6
3 to 4 years	23	10 to 11 years	6
4 to 5 years	43	11 to 12 years	6
5 to 6 years	42	12 to 13 years	3
6 to 7 years	24	> 13 years	5

Panel D: Distribution of Firms' Four-Digit SICs

Four-Digit SIC	SIC Industry Description	# Firms	% Firms
2834	Pharmaceutical preparations	65	32
8731	Commercial, physical, and biological research	50	24
2836	Biological products (except diagnostic substances)	20	11
3845	Electromedical and electrotherapeutic apparatus	12	6
2835	*In vitro* and *in vivo* diagnostic substances	12	6
All others	Various (representing 17 four-digit SICs)	45	20

(*continued on next page*)

TABLE 1 (continued)

Panel E: Location of Firm's Headquarters at Time of IPO Filing

State	# Firms	% Firms
California	94	46
Massachusetts	28	14
Pennsylvania	11	5
All others (20 states)	71	35

The venture capital market sample is the set of U.S. biotech firms in Recombinant Capital's (Recap) pre-IPO valuation database that met three criteria: The firm filed to go public; its filing was online at http://www.sec.gov; and the valuation date was both less than 12 months after any of the firm's pre-IPO filing fiscal year-ends and was the first financing event after that fiscal year-end. The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. The public equity market sample is the set of Recap firms that had an IPO filing document online and also completed their IPO.

statements—would be usable. Financings that are not explicitly coded by Recap as being Series A, Series B, etc. are excluded from Table 2,[14] as is the IPO filing or IPO offering. The reason for excluding the IPO itself is that a substantial body of research concludes that the equity values established at IPOs are associated with a number of anomalies (e.g., Shiller 1990; Ritter 1991; Ritter and Loughran 1995, 2002; Purnanandam and Swaminathan 2004). There is also cause for concern as to whether equity values computed using the price range in the prospectus or the final offer price are arrived at in an arms-length manner, given the collusive behavior that has been attributed to underwriters (e.g., Ritter 2000).

Table 2 shows that pre-money valuations increase as firms mature within the venture capital market, although at a decreasing rate. Valuations are much higher in the public market, partly because firms raise large amounts of cash at their IPO, and partly because going public may remove an illiquidity discount of 15 percent to 40 percent (Houlihan Valuation Advisors/VentureOne 1998).

A pattern of increasing median values is observed for almost all financial statement items except net income, which decreases steadily because biotech firms have small revenues but large expenditures on SG&A and R&D. Book-to-market ratios remain low during the time that firms are private, venture-backed entities. Of the 481 observations in the underlying pre-money valuation database, only 1 percent have a book-to-market ratio greater than 1. Consistent with firms' investment opportunity sets becoming smaller as firms mature, the median ratio of assets-in-place to pre-money equity value increases from 0.10 at Series A rounds, to 0.16 at Series C rounds, to 0.21 at Series F rounds, to 0.40 at the end of the first year after an IPO, and to 0.46 at the end of the third year after an IPO. This supports a crucial assumption underlying the hypothesis that the value relevance of financial statements will rise as firms mature.

Table 2 also reports the medians of several nonfinancial information items. For example, the age of the firm increases as the funding Series increases.[15] The median gap between

---

[14] Such financings include investments of common equity, private placements, and debt-related financing such as bridge and convertible notes. These can occur anywhere in the firm's pre-IPO life, whereas Series C financings never precede Series B financings, etc.

[15] One exception is that the median age of Series B financings is slightly less than that of Series A financings. Closer examination of this anomaly indicates that this was not due to data errors, but arose by chance because the procedure by which valuations and financial statement data are combined does not guarantee that age will be increasing in the Series level. It is also the case that sometimes a firm's first financing is labeled as Series B.

TABLE 2
Medians of Key Variables at Venture Capital Funding Rounds and Public Market Valuation Dates
(204 U.S. Biotech Firms, 1992–2003)

	Venture Capital Market Funding Round						Public Equity Market Year		
	Series A	Series B	Series C	Series D	Series E	Series ≥ F	IPO+1	IPO+2	IPO+3
Valuation Data									
Pre-money firm equity value ($ mil.)	6.3	15	27	41	53	55	143	142	111
Round-to-round raw return (mean)	NA	124%	92%	53%	37%	24%	2%	13%	17%
Level of AMEX public biotech index	137	137	151	144	169	161	190	336	336
Level of Sand Hill venture biotech index	494	551	726	844	811	827	1146	1778	1778
Financial Statement Data (fiscal year preceding the valuation date; $ mil.)									
Cash	0.2	1.5	2.6	3.6	5.6	9.3	42	38	33
Noncash assets	0.4	0.7	1.2	1.9	2.6	2.5	8.9	14	18
Long-term debt	0.0	0.1	0.2	0.5	0.5	0.7	0.6	0.6	1.3
Shareholder equity	−0.2	1.3	2.3	3.1	5.7	7.8	49	47	41
Revenue	0.03	0.0	0.05	0.2	1.2	1.9	5.9	8.2	9.6
SG&A costs	0.6	0.8	1.3	1.7	2.1	2.4	5.3	7.8	9.0
R&D expense	0.9	1.3	2.6	4.5	5.5	6.0	12	16	17
Net income	−1.0	−1.8	−3.5	−5.4	−6.9	−6.6	−12	−17	−20
Money raised at financing round	5.6	6.0	9.2	8.6	9.8	11	NA	NA	NA
Stock option dilution	6.9%	6.6%	7.2%	11%	11%	12%	9.5%	11%	13%
Total assets ÷ pre-money equity value	0.10	0.17	0.16	0.17	0.20	0.21	0.40	0.44	0.46

(*continued on next page*)

TABLE 2 (continued)

	Venture Capital Market Funding Round						Public Equity Market Year		
	Series A	**Series B**	**Series C**	**Series D**	**Series E**	**Series ≥ F**	**IPO+1**	**IPO+2**	**IPO+3**
Nonfinancial Statement Information									
Firm age at valuation date (yrs)	2.6	2.5	3.2	4.0	4.8	6.9	6.2	7.2	8.2
Prior fiscal year-end to valuation date (yrs)	0.18	0.34	0.34	0.34	0.26	0.34	0.25	0.25	0.25
Equity dilution created at financing date	46%	28%	28%	19%	18%	17%	NA	NA	NA
Number of patents filed	0	1	2	3	6	8	9	11	12
Number of patents issued	0	0	0	0	0	1	2	4	7
Scope of patents filed	0	0.67	0.50	0.67	0.60	0.47	0.42	0.40	0.40
Number of strategic alliances	1	2	2	2	4	4	7	9	11
Rounds led by corporate investor	8%	12%	14%	20%	24%	24%	NA	NA	NA
# Observations	39	89	105	82	55	34	160	154	135

The full venture capital market sample is the set of U.S. biotech firms in Recombinant Capital's (Recap) pre-IPO valuation database that met three criteria: The firm filed to go public; its filing was online at http://www.sec.gov; and the valuation date was both less than 12 months after any of the firm's pre-IPO filing fiscal year-ends and was the first financing event after that fiscal year-end. The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. The table above excludes financing rounds not explicitly labeled as "Series A," "Series B," and the like. The public equity market sample is the set of Recap firms that had an IPO filing document online and also completed their IPO. IPO+*k* is three months after the *k*th fiscal year-end following the IPO. The pre-money value of a firm at a given financing round is the value of the firm's equity after the investment made by the venture investors less the dollars invested by the venture investors.

the date on which the pre-money valuation is established and the end of the preceding fiscal year is always four months or less and is smallest for Series A and largest for the IPO. This mitigates concerns that financial statement data are stale by the time the firm's current round of financing is undertaken and are staler the earlier the round. The median levels of the public market AMEX Biotech Index and the private market Sand Hill Biotech Index each rise steadily as firms mature. Equity dilution created at the current financing round declines from a high of 46 percent in Series A rounds to 17 percent in Series ≥ F rounds, while the probability that the lead investor is a corporation steadily increases. The median number of patents filed and issued and the number of the firm's strategic alliances all increase monotonically as firms mature.

V. REGRESSION METHODS

Measuring Value Relevance

In financial accounting research, financial statements are said to be value-relevant if they are associated with equity prices, values, or returns. Following prior studies, the value relevance of a set of financial statement data is measured using adjusted R^2. Specifically, I define the value relevance of financial statements as the adjusted R^2 that can be uniquely attributed to financial statements in a regression of equity values on financial statements and nonfinancial statement information.[16] The total adjusted R^2 from a regression of equity values on financial statements and nonfinancial statement information can be decomposed into three mutually exclusive parts—that which is unique to financial statements, that which is unique to nonfinancial statement information, and that which is common to both financial statements and nonfinancial statement information. The adjusted R^2 that is unique to financial statements is the total regression adjusted R^2 less the adjusted R^2 from a regression of equity values on nonfinancial information alone. Likewise, the adjusted R^2 that is unique to nonfinancial statement information is the total regression adjusted R^2 less the adjusted R^2 from a regression of equity values on financial statements alone. Mindful of the inferential dangers of comparing adjusted R^2s across samples (Brown et al. 1999; Gu 2002), I conduct auxiliary analyses to test whether such dangers are significant and conclude that they are not (see Section VI).

Regression Models

The main regression model that I use to assess the value relevance of financial statements and nonfinancial information in private and public equity markets is the log-linear model. This approach differs from the linear specifications applied in prior accounting research into value relevance, but it conforms to every analysis of venture capital valuations undertaken in the entrepreneurial finance literature (e.g., Lerner 1994a; Gompers and Lerner 1999, 2000a, 2000b; Nicholson et al. 2003; Seppä 2003). The log-linear model has also been employed extensively in economics, particularly for evaluating R&D (Hall 1993, 2000; Hall et al. 2001).[17] The advantages of the log-linear model include its flexibility in accommodating nonlinear relationships (such as those arising from real options that are prevalent in biotech firms' equity values, as well as the interactions between biased accounting and large expenditures on intangible assets) and the econometric robustness it provides in dealing with outliers in the underlying nonlogged data. The main drawback of the log-linear model is that it is difficult (although not impossible) to develop a model in which equity value is a log-linear function of financial statements or nonfinancial information from first principles.[18]

---

[16] Among others, Collins et al. (1997) and Barth et al. (1998) use incremental adjusted R^2 to measure value relevance.

[17] Kaplan and Ruback (1995) and Berger et al. (1996) are two infrequent instances of the use of log-linear valuation models in the public equity market.

The regression models that I estimate include financial statement data from only the most recent fiscal year immediately prior to the valuation date. Including earlier financial statement data and/or instruments for expected future net cash flows, particularly as the firm matures and more years of historical financial statement data exists, could expand such a simplistic view. However, past financial statement data are highly correlated with current financial statement data, and direct measures of expected future net cash flows such as analyst forecasts are unavailable for pre-IPO firms.[19]

All dollar regression variables, Z, are log-transformed according to $\log_e[Z + 1]$ where $Z \geq 0$ is in $000s. The following nondollar variables, Y, are log-transformed according to $\log_e[Y + 1]$: stock option dilution, the level of the AMEX Biotech Index, the level of the Sand Hill Biotech Index, new equity dilution at the financing date, the number of patents filed, the scope of patents filed, and the number of strategic alliances. Unreported statistics indicate that these log transformations dramatically reduce the skewness present in the raw data and achieve greater homoscedasticity in regression residuals relative to most models that are not log-linear (ranks of non-log-transformed data also achieve high homoscedasticity in regression residuals).

The possibility that regression residuals may be serially correlated and/or heteroscedastic because of calendar time clustering in equity valuations leads me to estimate the standard errors of regression parameters using Newey and West's (1987) covariance matrix and Hansen's (1982) GMM approach. This accommodates both serial dependence and conditional heteroscedasticity in the residuals and has a small loss of precision relative to OLS, should residuals in fact be homoscedastic and serially uncorrelated (Davidson and MacKinnon 1993). Inferences are unchanged if OLS standard errors are employed instead.

VI. EMPIRICAL TESTS

Regressions Where Different Financing Rounds Are Pooled within a Given Market

I first estimate regressions in which observations are pooled within a given market. This provides a test of the first and second hypotheses of the paper—namely, that financial statements will overall be value-relevant in the venture capital market, and that the partial correlations between equity values and key financial statement items will be nonzero and of the same sign in both markets. The results, found in Table 3, broadly support these hypotheses.

Table 3 shows that financial statements are similarly value-relevant in venture capital and public equity markets. Controlling for the effects of nonfinancial statement information, firms' equity values are reliably positively related to cash, noncash assets, and R&D expense and are reliably negatively related to long-term debt and stock option dilution. Untabulated tests show that of the coefficients on the eight financial statement variables, only those on cash and revenues reliably differ across markets. The strong significance of the coefficients on R&D and cash is consistent with the prominence that biotech venture capitalists accord to R&D and cash in their assessment of biotech firms' survivability and potential for success. The reliably negative coefficient on long-term debt dampens the potential criticism that the positive coefficients observed on cash, noncash assets, and R&D are illusory or overstated because such variables merely capture scale effects. Were that criticism valid,

---

[18] Ye and Finn (2000) motivate a log-linear model of firms' equity market values, book equity, and net income by demonstrating that if the log of 1 plus the return on equity follows an AR(1) process, and net dividends are zero, then equity market value emerges as a multiplicative function of book equity and net income. Beatty et al. (2001) derive a log-linear valuation model under the assumptions that stock valuation is first-degree homogenous in underlying valuation drivers, that accounting constructs measure such valuation drivers with multiplicative measurement error that is conditionally lognormal, and that the unconditional distribution of stock values is either diffuse or lognormal.

[19] Untabulated tests indicate that lagged financial statement data are rarely incrementally associated with current-period equity values.

TABLE 3

Log-Linear Regressions of U.S. Biotech Firms' Pre-Money Valuations (*PREMV*) on Financial Statement Data (*FS*) and Nonfinancial Statement Information (*NFS*)

1992–2003

$$LnPREMV_{ik} = \alpha + \sum_{b} \theta_b LnFS_{bik} + \sum_{c} \gamma_c LnNFS_{cik} + \varepsilon_{ik}$$

Independent Variables	Predicted Coefficient Sign	Venture Capital Market		Public Equity Market	
Financial Statement Data					
Ln (1 + cash)	+	0.11	(4.4)***	0.51	(6.9)***
Ln (1 + noncash assets)	+	0.18	(4.3)***	0.19	(3.3)***
Ln (1 + long-term debt)	−	−0.05	(−3.8)***	−0.03	(−2.3)***
Ln (1 + revenue)	+	−0.01	(−1.2)	0.04	(1.8)*
Ln (1 + cost of sales)	−	−0.01	(−0.7)	0.00	(0.1)
Ln (1 + SG&A)	+/−	0.05	(0.9)	−0.14	(−1.3)
Ln (1 + R&D expense)	+	0.15	(2.9)***	0.14	(1.5)
Ln (1 + stock option dilution)	−	−1.56	(−4.2)***	−2.50	(−3.3)***
Indicator: No stock option data	−	−0.26	(−3.1)***		
Nonfinancial Statement Information					
Series A indicator		−0.28	(−1.5)		
Series B indicator		−0.21	(−2.1)*		
Series C indicator		−0.15	(−1.7)		
Series D indicator		−0.10	(−1.0)		
Series E indicator		−0.16	(−1.6)		
Series ≥ F indicator		−0.14	(−1.1)		
Two years after IPO indicator				−0.07	(−0.8)
Three years after IPO indicator				−0.15	(−1.2)
Corporate lead investor indicator	+	0.21	(2.9)***		
Ln (AMEX public biotech index)	+	0.32	(3.6)***	0.19	(1.6)
Ln (Sand Hill venture biotech index)	+	0.17	(1.8)*		
Ln (1 + firm age in years)	−	−0.01	(−0.2)	−0.14	(−1.4)
Ln (1 + new equity dilution)	−	−1.21	(−11.7)***		
Ln (1 + number of patents filed)	+	0.07	(2.7)***	0.06	(1.5)
Ln (1 + scope of patents filed)	+	−0.21	(−3.0)***	−0.39	(−2.2)**
Ln (1 + number of strategic alliances)	+	0.09	(2.6)***	−0.05	(−0.8)
# observations		481	~	449	
# different firms		193		160	
OLS regression adj. R^2 [total]		76%	>	47%	
Adj. R^2 [unique to financial statement data]		13%	<	26%	
Adj. R^2 [unique to nonfinancial statement information]		18%	>	1%	
Adj. R^2 [common to both]		45%	>	20%	
Residual standard deviation		0.56	<	0.87	
Median absolute valuation error		32%	<	49%	

(*continued on next page*)

TABLE 3 (continued)

---

*,**,*** Denote coefficient estimates that are reliably significant at the 5 percent, 2.5 percent, and 1 percent levels, respectively, under a one-tailed test of the predicted sign on the coefficient. Where there is no sign prediction, significance levels are computed under a two-tailed test.

The venture capital market sample is the set of U.S. biotech firms in Recombinant Capital's (Recap) pre-IPO valuation database that met three criteria: The firm filed to go public; its filing was online at http://www.sec.gov; and the valuation date was both less than 12 months after any of the firm's pre-IPO filing fiscal year-ends and was the first financing event after that fiscal year-end. The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. The public equity market sample is the set of Recap firms that had an IPO filing document online and also completed their IPO.

All financial statement data are in $000s before log-transformation.

The dependent variable is $LnPREMV_{ik}$, the log of firm *i*'s pre-money equity value at financing round *k*. GMM asymptotic Z-statistics are in parentheses. The intercept is estimated but not reported. Rounds other than those formally labeled by Recombinant Capital as Series A, B, C, D, E, and ≥ F are reflected in the intercept.

The AMEX public biotech index is an equal-dollar-weighted index designed to measure the performance of a cross-section of companies in the biotechnology industry that are primarily involved in the use of biological processes to develop products or provide services. It was established with a benchmark value of 200 on October 18, 1991.

The Sand Hill venture biotech index is a proprietary, periodic, timely, value-weighted, and unbiased index of the venture capital value of pre-IPO privately funded biotech companies developed by Sand Hill Econometrics of Palo Alto, CA. It had a benchmark value of 100 in January 1989. See http://www.sandhillecon.com.

New equity dilution is defined as the number of shares issued in the private equity financing round divided by the number of pre-money shares outstanding. The number of patents filed is the number of patents that had been filed as of the valuation date and that were issued by the U.S. Patent Office by 11/1/03, according to http://www.uspto.gov. The scope of patents filed as of the valuation date is defined as the cumulative number of distinct four-digit International Patent Classification classes into which the firm's filed patents are assigned as of the valuation date, divided by the number of patents the firm had filed for as of the valuation date.

The number of strategic alliances at the valuation date is the sum of the firm's upstream and downstream alliances in place at that time. These data are obtained from Recombinant Capital's alliances database.

The median absolute valuation error is an economic measure of the regression goodness-of-fit. The valuation error is defined as the ratio of the actual equity value less the fitted equity value, divided by the fitted equity value.

the strongly positive univariate correlations between equity values and long-term debt would lead to a positive, not negative, estimated regression coefficient on long-term debt.

Two unexpected results emerge. The coefficient on revenue is not reliably positive in the venture capital market, and the coefficient on cost of sales is not reliably different from zero in either market. The former finding may reflect the transitory nature of young biotech firms' revenues, in that it takes many years for a biotech firm to develop a product and thereby generate recurring product revenue.[20] Consistent with this explanation, the coefficient on revenue in the public market—where recurring product revenue is more common—is reliably positive. Also, although SG&A expense is not reliably related to firms' equity values, this likely occurs because SG&A contains a mixing of period expenses that would be expected to be negatively related to equity value (such as rent on the firm's facilities), and costs that provide future benefits (such as salaries for senior management and scientific personnel) that would be expected to be positively related to equity value, particularly when the firm is young and intellectual capital is most vital.

Table 3 also provides evidence on the value relevance of nonfinancial statement information in the venture capital market. The signs and significance of coefficient estimates on almost all the nonfinancial statement variables in Table 3 agree with what prior work found. Pre-money valuations are increasing in the level of publicly traded biotech company values and the number of a firm's patents and alliances and are also higher if the lead investor is a corporation rather than venture fund. One discrepancy is the reliably negative coefficient on the scope of the firm's patents, which Lerner (1994a) finds to be reliably positive. This result is not due to controlling for financial statement data; untabulated results show that the coefficient is negative if no financial statement data are included in the regression. I speculate that the economics of patent protection may have changed over time such that it is now more valuable to hold a portfolio of narrowly scoped patents rather than broadly scoped ones.[21]

---

[20] Few biotech companies generate substantial amounts of recurring product revenues until their drugs have passed through the FDA hurdles, which can be as much as 15 years into the life of a biotech firm (see Figure 1). Instead, young biotech firms generate revenues from fixed-period collaborations, contracts, grants, licenses, and research.

The estimated coefficients on each of the two items of nonfinancial statement information that have not been previously examined in the entrepreneurial finance literature are reliably different from zero in the predicted directions. First, the coefficient on the SHB Index is reliably positive and not significantly different from the larger coefficient estimated on the AMEX Biotech Index. This is consistent with the proposition that the SHB Index is informative about the overall prospects of the biotechnology sector over and above that provided by the prices of publicly traded biotech companies. Second, the size and significance of the coefficient on equity dilution created by the current financing round strongly suggests that demand is far from being perfectly elastic in the venture capital arena. The latter conclusion is the same as that of Gompers and Lerner (2000b) and Seppä (2003). However, untabulated tests indicate that the proxy used in Gompers and Lerner (2000b) and Seppä (2003)—the total amount of funding committed to U.S. venture capital funds in the 12 months leading up to the valuation date—is insignificant when it is included in the regression alongside the equity dilution variable.

Robustness Tests for Pooled Regressions

Untabulated results indicate the findings reported in Table 3 are robust to several issues. First, some prior work has found higher pre-money valuations in western states, particularly California, most likely because of the value created by geographical concentrations of scientific capital, venture capital, or business networks (e.g., Gompers and Lerner 2000a; Seppä 2003).[22] This is not the case here. Other untabulated results indicate that coefficient estimates on indicators for California- and Massachusetts-headquartered firms and the major sub-industries in biotechnology are insignificant. Second, Gompers and Lerner (2000b) and Seppä (2003) find that pre-money valuations are higher when inflows of capital into venture funds over the 12 months leading up to the valuation date are higher. That is, firm valuations reflect less than perfectly elastic demand in that "too much money is chasing too few deals." The same result is found in this study, but only when the Sand Hill Biotech Index is replaced by the inflows of capital into venture funds over the prior 12 months. This is because the correlation between the level of the Sand Hill Biotech Index and inflows of capital into venture funds over the 12 months leading up to the valuation date is 0.95. If both variables are included in the regression, then the equity dilution created by the current financing round is found to be a superior proxy for the presence of less than perfectly elastic demand in that the estimated coefficient on equity dilution (asymptotic Z-statistic $= -11.7$, one-tailed p-value < 0.0001) far exceeds that on venture capital fund inflows

---

[21] Lerner's and my samples do not overlap in time, making this at least possible: Lerner's data period is 1973–1992, whereas mine is 1992–2003.

[22] The firms examined in Gompers and Lerner (2000a) and Seppä (2003) are not limited to biotech firms, but come from a broad cross-section of industries.

(asymptotic Z-statistic = −1.6, one-tailed p-value = 0.06). This is most likely because the equity dilution proxy is more firm-specific.

It is also the case that if the total number of strategic alliances is separated into upstream and downstream deals, then the coefficient estimate on each deal-type is reliably negative, but there is no reliable difference between the coefficient estimates. Finally, the same inferences as those in Table 3 emerge if a rank regression is estimated instead of a log-linear model. The only exception is that the estimated coefficient on the total number of alliances is no longer reliably positive at the 1 percent level under a one-tailed test, only at the 10 percent level.

Regressions Estimated on a Financing-Round-by-Financing-Round Basis

The third hypothesis of the paper is that financial statements will be value-irrelevant when a firm is founded but become increasingly value-relevant as the firm matures. The fourth is that nonfinancial statement information will become less value-relevant as a firm matures. The pooled regressions in Table 3 provide coarse confirmation of these hypotheses in that the results suggest that financial statements are, on average, less value-relevant in the venture capital market than they are in the public market, and *vice versa* for nonfinancial statement information. Specifically, the adjusted R^2 uniquely explained by financial statements is 13 percent in the venture capital market versus 26 percent in the public market, while the adjusted R^2 uniquely explained by nonfinancial information is 18 percent in the venture capital market and only 1 percent in the public market.

Table 4 provides more powerful tests of the dynamic behavior of the value relevance of financial statements and nonfinancial statement information as firms mature. Moving from left to right, Table 4 reports the results of round-by-round regressions that reflect increasing levels of firm maturity.[23] Compared to pooled regressions, round-by-round regressions mitigate the potential impacts of sample selection bias because they control for one factor associated with the likelihood of filing for an IPO, namely the firm's maturity (Cochrane 2004). Round-by-round regressions should also mitigate any time-varying effects of not explicitly controlling for risk.

The regressions estimated in Table 4 contain only the most significant of the variables included in Table 3's pooled regressions, and with cash and noncash assets combined into total assets. This constriction of the set of independent variables is applied because there are far fewer observations per round than in the pooled regressions, and because there are high correlations among some independent variables. Untabulated analysis indicates that the inferences drawn from the results in Table 4 are not sensitive to including a larger set of independent variables.

The key results in Table 4 are consistent with the third hypothesis. Financial statements are value-irrelevant at (and only at) the Series A round, the first formal financing point for

---

[23] In-depth conversations with venture capital investors, particularly those who specialize in biotechnology, indicated that they deliberately stage their financing so that a firm's financing Series stage is more indicative of its scientific and economic maturity than is its calendar age. For example, biotech venture capitalists will typically only invest in the first formal financing round—the Series A—if certain scientific and business elements are present or will immediately be made present as a result of the funding. For example, they might require the firm to have a scientifically innovative idea for a new drug candidate, one or more highly regarded scientists, and at least one well-respected and seasoned biotech business executive. To obtain further financing through a Series B round, and then a Series C round, etc., the firm would need to achieve a number of specified scientific and/or business milestones, with significant flexibility given as to how long those milestones take to achieve. Evidence consistent with this venture capital perspective on firm maturity is that if firm age is added to the private equity market pooled regression estimated in Table 3, its coefficient is insignificant (t-statistic = −0.2).

TABLE 4
Log-Linear Round-by-Round Regressions of Venture Capital and Public Equity Market Values (*PREMV*) on Financial Statement Data (*FS*) and Nonfinancial Statement Information (*NFS*)
1992–2003

$$LnPREMV_i = \alpha + \sum_d \theta_d LnFS_{di} + \sum_f \gamma_f LnNFS_{fi} + \varepsilon_i$$

	Predicted Sign on Coefficient	Venture Capital Market						Public Equity Market		
		Series A	Series B	Series C	Series D	Series E	Series ≥ F	IPO+1	IPO+2	IPO+3
Financial Statement Data										
Ln (1 + total assets)	+	0.04	0.25	0.40	0.24	0.36	0.96	0.70	0.87	0.62
		(0.3)	(2.5)***	(5.5)***	(3.0)***	(3.0)***	(7.1)***	(7.0)***	(7.2)***	(6.1)***
Ln (1 + R&D expense)	+	0.14	0.23	0.22	0.23	0.38	−0.62	0.13	0.14	0.18
		(0.9)	(2.1)**	(2.4)***	(3.1)***	(2.4)***	(−2.6)***	(1.4)	(1.7)*	(1.5)
Nonfinancial Statement Information										
Ln (AMEX public biotech index)	+	0.59	0.49	0.47	0.43	0.35	0.34	−0.01	−0.19	0.61
		(1.0)	(3.1)***	(5.3)***	(5.5)***	(3.3)***	(3.3)***	(−0.1)	(−0.9)	(2.5)***
Ln (1 + new equity dilution)	−	−1.11	−1.21	−1.20	−1.49	−1.07	−0.30			
		(−5.8)***	(−8.7)***	(−6.6)***	(−4.4)***	(−2.3)**	(−0.7)			
Ln (1 + number of patents filed)	+	0.37	−0.01	0.03	0.02	−0.02	0.10	0.04	0.05	0.18
		(1.8)*	(−0.1)	(0.5)	(0.5)	(−0.4)	(1.1)	(0.7)	(0.9)	(2.7)***
Ln (1 + number of strategic alliances)	+	−0.05	0.09	−0.00	−0.05	−0.07	0.14	−0.02	−0.06	−0.02
		(−0.2)	(1.1)	(−0.0)	(−0.8)	(−0.9)	(1.2)	(−0.3)	(−0.9)	(−0.2)
# Observations		39	89	105	82	55	34	160	154	135
OLS regression adj. R^2 [total]		45%	64%	75%	69%	58%	75%	45%	52%	41%
Adj. R^2 [unique to financial statements]		3%	18%	25%	19%	15%	34%	24%	35%	21%
Adj. R^2 [unique to nonfinl. information]		31%	26%	17%	16%	9%	5%	−1%	0%	3%
Median absolute valuation error		43%	37%	27%	24%	31%	22%	43%	47%	65%
Total assets + (equity value + long-term debt)		0.10	0.17	0.16	0.17	0.19	0.20	0.40	0.42	0.46
Pr > F(value relev. of finl. statements = 0)		0.17	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Pr > F(value relev. of nonfinl. st. info. = 0)		< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.08	0.84	0.33	0.03
Coef. of variation [total assets]		0.44	0.24	0.16	0.19	0.13	0.11	0.08	0.10	0.13
Residual standard deviation		0.80	0.62	0.47	0.47	0.55	0.46	0.71	0.81	1.09

(continued on next page)

TABLE 4 (continued)

---

*,**,*** Denote coefficient estimates that are reliably nonzero at the 5 percent, 2.5 percent, and 1 percent levels, respectively, under a one-tailed test of the coefficient sign prediction. Where there is no sign prediction, significance levels are computed under a two-tailed test.

The venture capital market sample is the set of U.S. biotech firms in Recombinant Capital's (Recap) pre-IPO valuation database that met three criteria: The firm filed to go public; its filing was online at http://www.sec.gov; and the valuation date was both less than 12 months after any of the firm's pre-IPO filing fiscal year-ends, and was the first financing event after that fiscal year-end. The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. The public equity market sample is the set of Recap firms that had an IPO filing document online and also completed their IPO.

All financial statement data are in $000s prior to being log-transformed.

For a given round, the dependent variable is $LnPREMV_i$, the log of firm i's pre-money equity value. GMM asymptotic Z-statistics are in parentheses. The intercept is estimated but not reported. Rounds other than those formally labeled by Recombinant Capital as Series A, B, C, D, E, and ≥ F are reflected in the intercept.

The AMEX public biotech index is an equal-dollar-weighted index designed to measure the performance of a cross-section of companies in the biotechnology industry that are primarily involved in the use of biological processes to develop products or provide services. It was established with a benchmark value of 200 on October 18, 1991.

New equity dilution is defined as the number of shares issued in the private equity financing round divided by the number of pre-money shares outstanding. The number of patents filed is the number of patents that had been filed as of the valuation date and that were issued by the U.S. Patent Office by 11/1/03, according to http://www.uspto.gov. The number of strategic alliances at the valuation date is the sum of the firm's upstream and downstream alliances in place at that time. These data were obtained from Recombinant Capital's alliances database.

venture capital investors. The p-value on the F-statistic testing the constraint that the coefficients on both total assets and R&D expense are zero at Series A rounds is an insignificant 0.17. The second key result is that the value relevance of financial statements reliably increases as firms mature. Figure 2 plots the adjusted R^2 uniquely explained by each type of information shown in Table 4 as a function of the round-by-round maturity of the firm, together with OLS trendlines. The one-tailed p-value on the positive slope of the OLS trendline in the R^2 uniquely attributable to financial statements is 0.03. Very similar results obtain if ranks are used instead of log-transformed data. At least for biotech firms, financial statements are value-irrelevant at the earliest venture capital financing point, but become steadily more value-relevant as firms mature.

Beyond this finding, several other results in Table 4 deserve mention. First, the median ratio of total assets to total firm value, a proxy for the relative size of assets-in-place versus future investment options (Smith and Watts 1992; Gaver and Gaver 1993; Adam and Goyal 2003), reliably increases as firms mature (the two-tailed p-value on the positive trendline slope is less than 0.001). In addition, the correlation between the median ratio of total assets to total firm value and the R^2 uniquely attributable to financial statements is 0.51 (one-tailed p-value is 0.09).

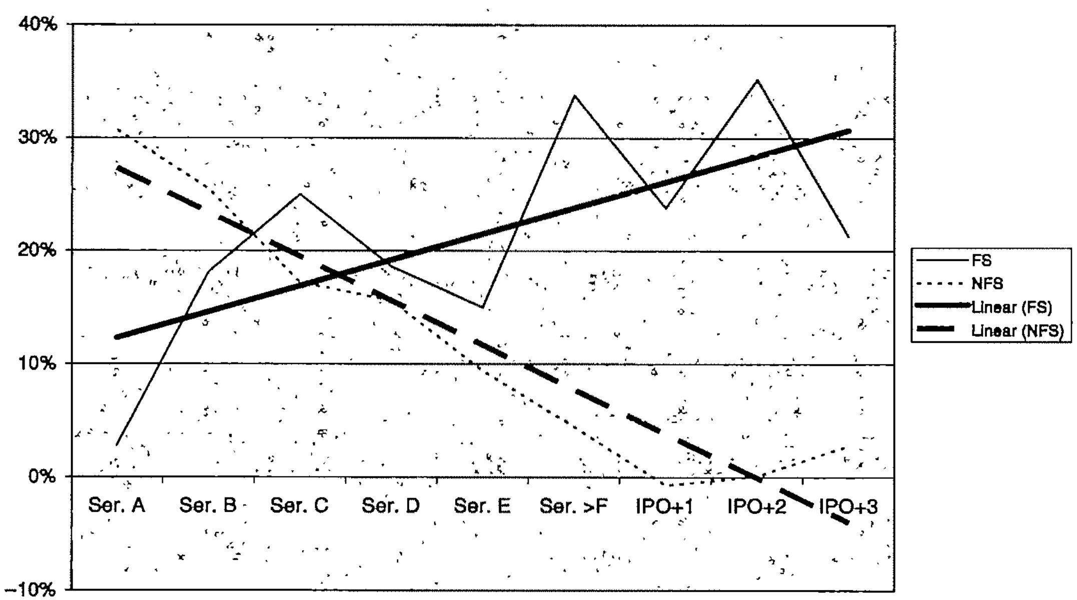
Second, consistent with the fourth hypothesis of the paper, the value relevance of nonfinancial information reliably decreases as firms mature. In Figure 2, the one-tailed p-value on the negative slope of the nonfinancial information unique R^2 trendline is 0.03. Virtually identical results obtain if ranks are used instead of log-transformed data. The time-series correlation between the value relevance of financial statement data and nonfinancial statement data is reliably negative (one-tailed p-value is 0.04), and that between the median ratio of total assets to total firm value and the value relevance of nonfinancial information is -0.82 (one-tailed p-value is 0.004). These indicate that in a dynamic sense, financial statements and nonfinancial statement data are informational substitutes, not complements, in valuation.

Third, obtaining an accurate view of the dynamic behavior of the value relevance of nonfinancial statement information hinges on controlling for the value relevance of financial statement data. Specifically, round-by-round regressions that exclude financial statements yield an essentially flat adjusted R^2 series of 43 percent, 46 percent, 50 percent, 51 percent, 44 percent, and 41 percent at Series A through Series $\geq$ F rounds, respectively. This contrasts with the starkly declining series of adjusted R^2s uniquely attributable to nonfinancial information of 31 percent, 26 percent, 17 percent, 16 percent, 9 percent, and 5 percent, respectively (Table 4), where the effects of financial statements are controlled for.

Fourth, as suggested by Table 3's results, Table 4 indicates that the lion's share of the value relevance of nonfinancial statement information in the venture capital market comes from the AMEX Biotech Index and the dilution created by the current financing round. On a round-by-round basis, the value relevance of patents and strategic alliances is very low. The same is true in public equity markets until the firm is three years or so beyond its IPO, at which point the AMEX Biotech Index and the number of patents each become value-relevant. The weak associations between equity values and the number of patents filed and strategic alliances entered into may reflect the simplicity or coarseness of these proxies. However, total assets are also a coarse and simple proxy for the economic value of assets-in-place.

Finally, the relations between value relevance and firm maturity are not artifacts due to scale or sampling variation. If this were so, then value relevance would be positively correlated with the coefficient of variation in candidate scale factors such as equity value or total assets (Brown et al. 1999, 2002). This is not the case; the Spearman correlations

FIGURE 2
Value Relevance of Financial Statement Data (*FS*) and Nonfinancial Statement Information (*NFS*) within and across the Pre-IPO Venture Capital and Post-IPO Public Equity Markets



Sample comprises an unbalanced panel of 204 U.S. biotech firms over the period 1992–2003. Value relevance ≡ VR(.) is the adjusted R^2 that is uniquely attributable to financial statement data (FS) or nonfinancial statement information (NFS), computed cross-sectionally at a particular pre-IPO venture capital financing round or post-IPO valuation point in time. For example, VR(FS) at Series B ("Ser. B") is the adjusted R^2 of a regression of venture capital pre-money firm equity valuations at Series B financing rounds on a set of financial statement and nonfinancial statement variables less the adjusted R^2 of a similar regression that only includes nonfinancial statement variables. The solid straight lines, denoted Linear (*FS*) and Linear (*NFS*), are the OLS slopes fitted to the time-series of VR(*FS*) and VR(*NFS*).

between the adjusted R^2 uniquely attributable to financial statements and the coefficients of variation of log-transformed equity value and log-transformed total assets are -0.39 and -0.64, respectively. Alternatively, if the pattern in value relevance were due to sample variances that were systematically increasing over time, then value relevance would be positively correlated with the standard deviation of the regression residuals (Gu 2002). This, too, is not the case. The correlation between the adjusted R^2 uniquely attributable to financial statements and the standard deviation of round-by-round regression residuals is -0.14, while the correlation between the adjusted R^2 uniquely attributable to nonfinancial statement information and the standard deviation of round-by-round regression residuals is -0.22.

The IPO

As noted in Section IV, this study has thus far deliberately excluded the IPO itself as a valuation point because a substantial body of literature has concluded that equity values set at and close to IPOs are associated with notable anomalies (e.g., Shiller 1990; Ritter 1991; Ritter and Loughran 1995, 2002; Ritter 2000; Purnanandam and Swaminathan 2004). Common to many explanations for IPO-related pricing anomalies is the proposition that IPOs are sold by sophisticated underwriters to unsophisticated retail investors—i.e., that the marginal investor is markedly less than fully rational during the early months of a newly public firm's life.

In the context of this study, the unsophisticated marginal investor proposition implies that the value relevance of financial statement data (nonfinancial statement information) at the IPO filing will be lower (higher) than that implied by the maturity-based trends reported in Table 4 and Figure 2. The rationale is that unsophisticated investors are likely to place too much weight on the value relevance of nonfinancial statement information, particularly herd-related variables such as the "hotness" of the IPO market, and therefore too little weight on the true value relevance of financial statement data.

Table 5 reports the results of testing these predictions. Table 5 is structured in the same manner as Table 3, except that in between the columns of results for the venture capital and public equity markets are the results of estimating the same pre-money equity valuation regression at the IPO filing. The IPO filing results are consistent with the predictions made by the unsophisticated investor proposition. Specifically, at the IPO filing, the adjusted R^2s uniquely attributable to financial statements and nonfinancial information are 2 percent and 28 percent, respectively. These are almost exactly the reverse of what would be expected based on the trendlines in unique adjusted R^2s shown in Figure 2. Among financial statement data, only noncash assets are reliably associated with firms' pre-money IPO equity values. In contrast, among nonfinancial items of information, the AMEX Biotech Index and number of patents filed are reliably positively related to pre-money IPO equity values, and the new equity dilution created at the IPO is reliably negatively related. Both the AMEX Biotech Index and the new equity dilution created at the IPO can be interpreted in an IPO context as measures of herd-related forces such as the "hotness" of the IPO market. It therefore appears that the value relevance of both financial statement data and nonfinancial statement information to some extent follows a "sawtooth" pattern when the IPO is included in the maturity timeline of U.S. biotech firms.

Caveats and Limitations

Beyond the straightforward concern as to how well the results reported here for venture-backed biotech firms generalize to the population of all pre-IPO firms, this paper has two main limitations: differences between the venture capital and public equity markets that have not been controlled for, and the uncertain impacts of selection bias.

As detailed in Section II, the venture capital market differs from the public equity market along many dimensions, including regulation, liquidity, price-setting mechanisms, and degree of information asymmetry between managers and investors. Such differences may affect the value relevance of financial statements in ways that have not been taken into account in this paper. However, it is difficult to pinpoint in what ways and how strongly

TABLE 5

A Comparison of Log-Linear Value Relevance Regressions for U.S. Biotech Firms in the Venture Capital Market, at the IPO Filing, and in the Public Equity Market 1992–2003

$$LnPREMV_{ik} = \alpha + \sum_{b} \theta_b LnFS_{bik} + \sum_{c} \gamma_c LnNFS_{cik} + \varepsilon_{ik}$$

Independent Variables	Venture Capital Market		At the IPO Filing		Public Equity Market	
Financial Statement Data						
Ln (1 + cash)	0.11	(4.4)***	0.06	(1.0)	0.51	(6.9)***
Ln (1 + noncash assets)	0.18	(4.3)***	0.16	(2.3)**	0.19	(3.3)***
Ln (1 + long-term debt)	−0.05	(−3.8)***	−0.02	(−1.5)	−0.03	(−2.3)***
Ln (1 + revenue)	−0.01	(−1.2)	−0.02	(−1.5)	0.04	(1.8)*
Ln (1 + cost of sales)	−0.01	(−0.7)	−0.01	(−0.8)	0.00	(0.1)
Ln (1 + SG&A)	0.05	(0.9)	−0.11	(−1.0)	−0.14	(−1.3)
Ln (1 + R&D expense)	0.15	(2.9)***	0.04	(0.6)	0.14	(1.5)
Ln (1 + stock option dilution)	−1.56	(−4.2)***	0.10	(0.1)	−2.50	(−3.3)***
Indicator: No stock option data	−0.26	(−3.1)***				
Nonfinancial Statement Information						
Series A indicator	−0.28	(−1.5)				
Series B indicator	−0.21	(−2.1)*				
Series C indicator	−0.15	(−1.7)				
Series D indicator	−0.10	(−1.0)				
Series E indicator	−0.16	(−1.6)				
Series ≥ F indicator	−0.14	(−1.1)				
Two years after IPO indicator					−0.07	(−0.8)
Three years after IPO indicator					−0.15	(−1.2)
Corporate lead investor indicator	0.21	(2.9)***				
Ln (AMEX public biotech index)	0.32	(3.6)***	0.54	(2.7)***	0.19	(1.6)
Ln (Sand Hill venture biotech index)	0.17	(1.8)*	0.19	(0.7)		
Ln (1 + firm age in years)	−0.01	(−0.2)	0.08	(1.0)	−0.14	(−1.4)
Ln (1 + new equity dilution)	−1.21	(−11.7)***	−2.71	(−3.9)***		
Ln (1 + number of patents filed)	0.07	(2.7)***	0.13	(3.1)***	0.06	(1.5)
Ln (1 + scope of patents filed)	−0.21	(−3.0)***	−0.08	(−0.5)	−0.39	(−2.2)**
Ln (1 + number of strategic alliances)	0.09	(2.6)***	0.09	(1.5)	−0.05	(−0.8)
# Observations	481		103		449	
# Different firms	193		103		160	
OLS regression adj. R^2 [total]	76%		72%		47%	
Adj. R^2 [unique to *FS* data]	13%		2%		26%	
Adj. R^2 [unique to *NFS* information]	18%		28%		1%	
Adj. R^2 [common to both]	45%		42%		20%	
Residual standard deviation	0.56		0.44		0.87	
Median absolute valuation error	32%		26%		49%	

(*continued on next page*)

TABLE 5 (continued)

---

*,**,*** Denote coefficient estimates that are reliably significant at the 5 percent, 2.5 percent and 1 percent levels, respectively, under a one-tailed test of the sign on the coefficient predicted in Table 3. Where there is no sign prediction, significance levels are computed under a two-tailed test.

The venture capital market sample is the set of U.S. biotech firms in Recombinant Capital's (Recap) pre-IPO valuation database that met three criteria: The firm filed to go public; its filing was online at http://www.sec.gov; and the valuation date was both less than 12 months after any of the firm's pre-IPO filing fiscal year-ends and was the first financing event after that fiscal year-end. The valuation date is the date the terms of the financing round were agreed on by the firm and the venture investors. The IPO filing sample is the set of Recap firms that had an IPO filing document online. The public equity market sample is the set of Recap firms that had an IPO filing document online and also completed their IPO. All financial statement data are in $000s before log-transformation.

The dependent variable is $LnPREMV_{ik}$, the log of firm *i*'s pre-money equity value at financing round *k*. GMM asymptotic Z-statistics are in parentheses. The intercept is estimated but not reported. Rounds other than those formally labeled by Recap as Series A, B, C, D, E, and ≥F are reflected in the intercept.

The AMEX public biotech index is an equal-dollar-weighted index designed to measure the performance of a cross-section of companies in the biotechnology industry that are primarily involved in the use of biological processes to develop products or provide services. It was established with a benchmark value of 200 on October 18, 1991.

The Sand Hill venture biotech index is a proprietary, periodic, timely, value-weighted, and unbiased index of the venture capital value of pre-IPO privately funded biotech companies developed by Sand Hill Econometrics of Palo Alto, CA. It had a benchmark value of 100 in January 1989. See http://www.sandhillecon.com.

New equity dilution is defined as the number of shares issued in the private equity financing round divided by the number of pre-money shares outstanding. The number of patents filed is the number of patents that had been filed as of the valuation date and that were issued by the U.S. Patent Office by 11/1/03, according to http://www.uspto.gov. The scope of patents filed as of the valuation date is defined as the cumulative number of distinct four-digit International Patent Classification classes into which the firm's filed patents are assigned as of the valuation date, divided by the number of patents the firm had filed for as of the valuation date.

The number of strategic alliances at the valuation date is the sum of the firm's upstream and downstream alliances in place at that time. These data are obtained from Recap's alliances database. The median absolute valuation error is an economic measure of the regression goodness-of-fit. The valuation error is defined as the ratio of the actual equity value less the fitted equity value, divided by the fitted equity value.

such differences will affect the average levels of the value relevance of financial statement data and nonfinancial statement information in the two markets or the dynamic behavior of value relevance as firms mature.

With regard to biases that may arise from sample selection criteria, restricting the sample to firms that were successful enough to file for an IPO yields benefits in the form of rich financial statement data, but also carries risks. Specifically, the coefficients on the determinants of pre-money valuations may be biased if financial statements or nonfinancial information are more value-relevant for successful firms. Any bias may also be more severe the earlier the financing round to which the pre-money valuation applies.

It may be that any inferential biases are small. First, the firms used in this study consist of a truncated sample in that neither valuations nor valuation determinants are observed for firms that do not file to go public (because they are bought out, terminate, or continue in existence without new private equity funding or an IPO filing). The standard result of truncating on variable y, when the true relation between y and x is linear, is to bias toward zero the slope coefficient in a regression of observed y on observed x. This suggests that the slope coefficients on the determinants of valuations may be biased toward, not away from, zero. Second, the pooled regressions include financing round indicators in pre-money valuation regressions. The presence of these indicators should mitigate bias on slope coefficients that would otherwise arise from more mature rounds having larger mean values of the pre-money valuation dependent variable (Maddala 1977). Finally, in their study of the impact of fund inflows on venture-backed firms' pre-money valuations, Gompers and

Lerner (2000b) find that using Heckman's (1976) two-stage method to correct for selection bias has little impact on the magnitude or the significance of their independent variables.[24]

VII. CONCLUSIONS

This study contributes to the value relevance literatures in accounting and entrepreneurial finance by documenting the level, nature, and dynamics of the value relevance of financial statements in the venture capital subsector of the private equity market, and over the previously unexamined early stages of the firm life-cycle. Using a sample of U.S. biotechnology firms during 1992–2003, I found that financial statements are, on average, value-relevant in the venture capital market, despite the fact that such markets are different from the public market in numerous economically significant ways. Financial statements are also similarly associated with equity values in economically sensible ways in each market: equity values are positively related to firms' cash balances, noncash assets, and R&D expense, and are negatively related to their long-term debt and stock option dilution. This indicates that, as intended by the FASB in setting up the Conceptual Statements and Issued Standards, GAAP successfully provides information that is useful to investors outside of the public equity market. This is true even for firms such as young biotech companies whose equity values are dominated by expectations of profitable future investments rather than assets-in-place.

I also showed that financial statement data are less value-relevant in the venture capital market than is nonfinancial statement information, and that the opposite holds true in the public equity market. Specifically, financial statements and nonfinancial statement information for biotech companies are in a dynamic sense information substitutes, not complements. As firms mature, the value relevance of their financial statements reliably increases, while the value relevance of nonfinancial statement data reliably decreases. I provided evidence that this divergence occurs because as firms mature, the importance of assets-in-place increases relative to that of investment opportunities, and financial statement data are better than nonfinancial information at reflecting the value stemming from assets-in-place. The only exception to these patterns occurs at the IPO, where the value relevance of financial statement data and nonfinancial statement information reverse in importance. This "sawtooth" in the maturity trends in value relevance appears to occur because the marginal investor in an IPO is markedly less sophisticated than either venture capital investors in the pre-IPO venture capital market, or investors in seasoned post-IPO firms.

Finally, beyond documenting the strong relations between financial statement data and equity value in the venture capital market, this paper extends the literature in entrepreneurial finance by combining into one analysis most of the nonfinancial statement information variables examined in prior work, and also by identifying two new value-relevant nonfinancial factors. In particular, the equity dilution created by the current round is shown to be a very significant factor in explaining variation in firms' pre-money valuations, indicating that the elasticity of demand in the venture capital market is markedly less than infinite. Given this, an appealing topic for future research to explore might be to identify the determinants of the elasticity of demand, and whether and why financial statements are included in the set of those determinants. It would also be fruitful to determine whether the results found for biotech firms generalize to other technology firms, as well as non-technology companies, which sometimes obtain venture capital funding.

---

[24] Gompers and Lerner (2000b) first estimate the probability that VentureOne, their data provider, was able to obtain information about the valuation in the financing round, and then seek to explain the determinants of the valuation after adjusting for that probability.

REFERENCES

Adam, T., and V. K. Goyal. 2003. The investment opportunity set and its proxy variables: Theory and evidence. Working paper, Hong Kong University of Science & Technology.

Amir, E., and B. Lev. 1996. Value relevance of nonfinancial information: The wireless communications industry. *Journal of Accounting & Economics* 22: 3–30.

Barth, M. E., W. H. Beaver, and W. R. Landsman. 1998. Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting & Economics* 25: 1–34.

———, ———, and ———. 2001. The relevance of the value relevance literature for financial accounting standard setting: Another view. *Journal of Accounting & Economics* 31: 77–104.

Beatty, R., S. Riffe, and R. Thompson. 2001. The log-linear form of stock value models based on accounting information. Working paper, Southern Methodist University.

Berger, P. G., E. Ofek, and I. Swary. 1996. Investor valuation of the abandonment option. *Journal of Financial Economics* 42: 257–287.

Black, E. L. 1998. Life cycle impacts on the incremental value relevance of earnings and cash flow measures. *Journal of Financial Statement Analysis* Fall: 40–56.

Brav, A., and P. A. Gompers. 1997. Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture- and nonventure-capital-backed companies. *Journal of Finance* 52: 1791–1822.

Brown, S., K. Lo, and T. Lys. 1999. Use of R-squared in accounting research: Measuring changes in value relevance over the last four decades. *Journal of Accounting & Economics* 28: 83–115.

———, ———, and ———. 2002. Erratum to "Use of R-squared in accounting research: Measuring changes in value relevance over the last four decades." *Journal of Accounting & Economics* 33: 141.

Cochrane, J. H. 2004. The risk and return of venture capital. Working paper, University of Chicago.

Collins, D. W., E. L. Maydew, and I. S. Weiss. 1997. Changes in the value relevance of earnings and book values over the past forty years. *Journal of Accounting & Economics* 24: 39–67.

Darby, M. R., Q. Liu, and L. G. Zucker. 1999. Stars and stakes: The effect of intellectual human capital on the level and variability of high-tech firms' market values. Working paper, University of California, Los Angeles.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford, U.K.: Oxford University Press.

Deng, Z., B. Lev, and F. Narin. 1999. Science and technology as predictors of stock performance. *Financial Analysts Journal* 55 (5): 20–32.

Ely, K., P. Simko, and L. G. Thomas. 2003. The usefulness of biotechnology firms' drug development status in the evaluation of research and development costs. *Journal of Accounting, Auditing and Finance* 18: 163–196.

Fenn, G. W., N. Liang, and S. Prowse. 1995. *The Economics of the Private Equity Market*. Washington, D.C.: Board of Governors of the Federal Reserve System.

Financial Accounting Standards Board (FASB). 1978. Statement of Financial Accounting Concepts No. 1. Norwalk, CT: FASB.

Gaver, J. J., and K. M. Gaver. 1993. Additional evidence on the association between the investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Accounting & Economics* 16: 125–160.

Gompers, P. A. 1995. Optimal investment, monitoring, and the staging of venture capital. *Journal of Finance* 50: 1461–1490.

———, and J. Lerner. 1999. An analysis of compensation in the U.S. venture capital partnership. *Journal of Financial Economics* 51: 3–44.

———, and ———. 2000a. *The Venture Capital Cycle*. Cambridge, MA: MIT Press.

———, and ———. 2000b. Money chasing deals? The impact of fund inflows on the valuation of private equity investments. *Journal of Financial Economics* 55: 281–325.

Gu, Z. 2002. Cross-sample incomparability of R^2s and additional evidence on value relevance changes over time. Working paper, Carnegie Mellon University.

Guo, R., B. Lev, and N. Zhou. 2004. Competitive costs of disclosure by biotech IPOs. *Journal of Accounting Research* 42 (2): 319–355.

Hall, B. H. 1993. The stock market's valuation of research and development investment during the 1980s. *American Economic Review* 83: 259–264.

———. 2000. Innovation and market value. In *Productivity, Innovation and Economic Performance*, edited by R. Barrell, G. Mason, and M. O'Mahoney. Cambridge, U.K.: Cambridge University Press.

———, A. Jaffe, and M. Trajtenberg. 2001. Market value and patent citations: A first look. NBER Working Paper No. 7741. Cambridge, MA: National Bureau of Economic Research.

Hand, J. R. M. 2003. Profits, losses and the nonlinear pricing of Internet stocks. In *Intangible Assets: Values, Measures, and Risks*, edited by J. R. M. Hand, and B. Lev. Oxford, U.K.: Oxford University Press.

———. 2004. The market valuation of biotechnology firms and biotechnology R&D. In *Venture Capital Contracting and the Valuation of High-Technology Firms*, edited by J. McCahery, and L. Renneboog. Oxford, U.K.: Oxford University Press.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029–1054.

Heckman, J. J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5: 475–492.

Houlihan Valuation Advisors/VentureOne. 1998. The pricing of successful venture-backed high-tech and life-science companies. *Journal of Business Venturing* 13: 331–351.

Ittner, C., and D. Larcker. 1998. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 6: 205–238.

Joos, P. 2002. Explaining cross-sectional differences in market-to-book ratios in the pharmaceutical industry. Working paper, University of Rochester.

Kaplan, S. N., and R. S. Ruback. 1995. The valuation of cash flow forecasts: An empirical analysis. *Journal of Finance* 50: 1059–1093.

———, and P. Stromberg. 2003. Financial contracting meets the real world: An empirical analysis of venture capital contracts. *Review of Economics Studies* 70: 281–315.

———, and ———. 2004. Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *Journal of Finance* 59 (5): 2177–2210.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting & Economics* 31: 105–231.

Lerner, J. 1994a. The importance of patent scope: An empirical analysis. *Rand Journal of Economics* 25: 319–333.

———. 1994b. Venture capitalists and the decision to go public. *Journal of Financial Economics* 35: 293–316.

Lev, B., and P. Zarowin. 1999. The boundaries of financial reporting and how to extend them. *Journal of Accounting Research* 37: 353–385.

———. 2001. *Intangibles: Management, Measurement, and Reporting*. Washington, D.C.: Brookings Institution Press.

Maddala, G. S. 1977. *Econometrics*. New York, NY: McGraw-Hill.

Miller, M. H., and F. Modigliani. 1961. Dividend policy, growth and the valuation of shares. *Journal of Business* 34: 411–433.

Myers, S. 1977. Determinants of corporate borrowing. *Journal of Financial Economics* 5: 147–175.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Nicholson, S., P. M. Danzon, and J. McCullough. 2003. Biotech-pharmaceutical alliances as a signal of asset and firm quality. NBER Working Paper 9007. Cambridge, MA: National Bureau of Economic Research.

Purnanandam, A. K., and B. Swaminathan. 2004. Are IPOs really underpriced? *Review of Financial Studies* 17 (3): 811–848.

Ritter, J. R. 1991. The long run performance of initial public offerings. *Journal of Finance* 46: 3–28.

———, and T. Loughran. 1995. The new issues puzzle. *Journal of Finance* 50: 23–51.

———. 2000. The seven percent solution. *Journal of Finance* 55: 1105–1131.

———, and T. Loughran. 2002. Why don't issuers get upset about leaving money on the table in IPOs? *Review of Financial Studies* 15: 413–443.

Robinson, D. T., and T. Stuart. 2002. Contractual, relational, and equity-based control in biotech strategic alliances. Working paper, Duke University.

Sahlman, W. A. 1990. The structure and governance of venture capital organizations. *Journal of Financial Economics* 27: 473–521.

Seppä, T. J. 2003. Essays on the valuation and syndication of venture capital investments. Ph.D. dissertation, Helsinki University of Technology.

Shiller, R. 1990. Speculative prices and popular models. *Journal of Economic Perspectives* 4: 55–65.

Skinner, D. 1993. The investment opportunity set and accounting procedure choice: Preliminary evidence. *Journal of Accounting & Economics* 16: 407–445.

Smith, C. W., and R. L. Watts. 1992. The investment opportunity set and corporate financing, dividends, and compensation policies. *Journal of Financial Economics* 32: 262–292.

Smith, K. G., T. R. Mitchell, and C. E. Summer. 1985. Top level management priorities in different stages of the organizational life cycle. *Academy of Management Journal* 28: 729–820.

Stuart, T. E., H. Hoang, and R. C. Hybels. 1999. Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly* 44: 315–349.

Trueman, B., F. Wong, and X-J. Zhang. 2000. The eyeballs have it: Searching for the value in Internet stocks. *Journal of Accounting Research* 38 (Supplement): 137–162.

Wong, A. 2002. Angel finance: The other venture capital. Working paper, University of Chicago.

Wright, M., and K. Robbie. 1998. Venture capital and private equity: A review and synthesis. *Journal of Business Finance & Accounting* 25 (5) & (6): 521–570.

Ye, J., and M. Finn. 2000. Nonlinear accounting-based equity valuation models. Working paper, Baruch College.

Zhang, X-J. 2001. Accounting conservatism and the analysis of line items in earnings forecasting and equity valuation. Working paper, University of California, Berkeley.

Zucker, L. G., and M. R. Darby. 1996. Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences* 93 (23): 12709–12716.

———, ———, and M. B. Brewer. 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. *American Economic Review* 88 (1): 290–306.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 649–675

Audit Committee Quality and Internal Control: An Empirical Analysis

Jayanthi Krishnan
Temple University

ABSTRACT: I examine the association between audit committee quality and the quality of corporate internal control. While information on the quality of internal control is not generally available, companies changing auditors are required to disclose any internal control problems that were pointed out by their predecessor auditors. The empirical results are based on a comparison of companies disclosing such internal control problems with a control sample of companies changing auditors but not disclosing internal control problems. Audit committee quality is measured in three dimensions: its size, its independence, and its expertise. The internal control problems are observed at two levels of increasing seriousness: reportable conditions and material weaknesses. The sample time period precedes the effective dates of recent policy changes regarding audit committees. The results indicate that independent audit committees and audit committees with financial expertise are significantly less likely to be associated with the incidence of internal control problems. This is true for both levels of internal control problems. The results are consistent with recent policy emphasis on audit committee independence and expertise.

Keywords: *audit committee; internal control; reportable condition; material weakness.*

Data Availability: *Data used in this study are from public sources.*

I. INTRODUCTION

The Sarbanes-Oxley Act of 2002 and the Securities and Exchange Commission (SEC) rules introduced to carry out the directions of the Act indicate that improving corporate internal control is an important policy initiative (U.S. House of Representatives 2002). The SEC now requires CEOs and CFOs to certify annually and quarterly about the effectiveness of their corporate internal control (SEC 2002). Further, the SEC has announced rules for the implementation of Section 404 of the Sarbanes-Oxley Act, requiring every registrant to include an internal control report in its annual report, and its independent auditor to issue an "attestation report on management's assessment of the company's internal control over financial reporting" (SEC 2003a, 1).

I appreciate the helpful comments of two anonymous reviewers, Dana Hermanson (discussant at the 2001 AAA Annual Meeting), Jagan Krishnan, Roland Lipka, Eric Press, Sri Ramamoorti, Heibatollah Sami, Arnie Wright, and seminar participants at the 2001 AAA Annual Meeting and 2003 Mid-Atlantic Regional Meeting. Financial support was provided by the Fox School of Business and Management, Temple University, and Temple University's Summer Research Fellowship.

Editor's note: This paper was accepted by Mark Nelson, Guest Editor.

Submitted June 2003
Accepted August 2004

The emphasis on good internal control of course arises because it is considered to be an important factor in achieving good quality financial reporting. Among the numerous criticisms leveled at Enron Corporation regarding its financial reporting was the charge of a failure in its internal control (Verschoor 2002). It is therefore important to examine monitoring mechanisms corporations can use to ensure the effectiveness of their internal control. Establishing and maintaining proper internal control is the responsibility of management. One monitor of an entity's internal control is its audit committee.[1] Although not mandated by SEC rules, the accounting profession and policy makers have maintained that a primary function of the audit committee should be the oversight of internal control.[2] Anecdotal evidence (e.g., disclosures in proxy filings) and academic research suggest audit committees do view monitoring internal controls as one of their functions. Carcello et al. (2002, 294) report that 91 percent of a random sample of corporate audit committee charters list reviews of internal controls as one of the duties of the audit committee. Recently, the SEC (2003b) has mandated that all material written communications between a company's accountant and management be provided to the entity's audit committee, and this includes "reports on observations and recommendations on internal controls." The implication is that an entity's internal controls should be under the purview of its audit committee.

Because the audit committee is expected to monitor the entity's internal control, I posit a positive association between audit committee quality and internal control quality. What constitutes a good audit committee however is not clear. In 1999, the major exchanges introduced new requirements for the size, independence, and expertise of corporate audit committees with a view to improving financial reporting. In its endorsement of the exchanges' proposals to revise audit committee requirements, the SEC quoted the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees (BRC) to argue that, under the new requirements, audit committee directors would be "more likely to objectively evaluate the propriety of management's accounting, *internal control*, and financial reporting practices" (SEC 1999) (emphasis added). However, given the skepticism about the effectiveness of audit committee independence following recent disclosures about the composition of Enron's audit committee (which was independent as defined by the SEC), and concerns about the adequacy of the audit committee expertise requirement (Whitehead 2002), it is an open question whether the audit committee independence and expertise requirements can favorably affect the monitoring of internal control.

I examine the association between audit committee composition and the quality of corporate internal control, using data from a time period preceding the effective date (June 2001) for the SEC (1999) rules. Data on internal control quality is not generally available, which explains the absence of studies in this area. However, companies that change auditors are required to disclose any internal control problems that were pointed out by their predecessor auditors.[3] Using Form 8-K disclosures about internal control problems for a sample of companies that changed auditors, and a control sample of companies that changed auditors but did not report internal control problems, I find a negative association between the proportion of independent members on the audit committee and the existence of internal control problems. This result supports recent studies that have shown an association between audit committee independence and two aspects of the quality of financial reporting, the auditor's issuance of the going concern opinion for financially stressed companies (Carcello and Neal 2000), and the extent of earnings management as measured by discretionary accruals (Klein 2002b). I also find a negative association between the number of audit committee members with financial expertise and the existence of internal control problems. This finding suggests, in line with prior experimental studies regarding the association between audit committee expertise and financial reporting (e.g., McDaniel et al. 2002; DeZoort 1998), that audit committee financial expertise is associated with better controls.

---

[1] More completely, monitoring internal control is the result of actions by, and interaction between, management, the internal auditor (if present), the external auditor, and the audit committee. I focus on whether, holding other things constant, the audit committee has an impact on internal control quality.

[2] An early release by the SEC stressed the "importance of audit committees to enable board of directors to better fulfill their oversight responsibilities with respect to an issuer's accounting, financial reporting and control obligations" (SEC 1979, 427). The Treadway Commission's report recommended: "To be effective, audit committees should exercise vigilant and informed oversight of the financial reporting process, including the company's internal controls" (NCFFR 1987, 12).

The next section describes the hypothesis and research method. In Section III, I describe the data. This is followed in Sections IV and V by the empirical findings and additional tests. The last section contains a discussion of the results and my conclusions.

II. HYPOTHESIS AND RESEARCH METHOD

The quality of an entity's internal control is a function of the quality of its control environment (which includes the board of directors and the audit committee) and its other non-governance-related controls. The profession and policy makers have noted repeatedly that one of the *recommended* duties of the audit committee is to review the entity's internal control. For example, the BRC (1999, 38) recommends that "the audit committee should encourage procedures that promote accountability ... ensuring that management properly develops and adheres to a sound system of internal control."[4] Other policy-making bodies have, time and again, made similar recommendations.

While it seems clear that audit committees are *expected* to play a role in maintaining effective control, what happens in reality? There is some evidence that audit committees participate in the maintenance of control. Based on a survey of audit committee members, DeZoort (1997, 210) reports that "audit committee members perceive internal control evaluation as the most important audit committee oversight area." A study of a random sample of audit committee charters (Carcello et al. 2002) reveals that a majority of audit committees claim review of controls to be one of their duties. Examination of disclosures in proxy filings of companies in my sample also shows that in practice, corporate audit committees do engage in the examination of controls.

If audit committees are expected to oversee, and view themselves as overseers of, entities' internal control, then it seems plausible that a high-quality audit committee would be associated with a good-quality internal control. Other studies have found that the quality of the audit committee is associated with financial reporting outcomes (Carcello and Neal 2000; Klein 2002b). Thus I posit the following:

---

[3] I acknowledge that this sample is not representative of the population of firms, because auditor change firms have been shown in past work to be different from the general population of firms. (See Krishnan and Krishnan [1997] and Raghunandan and Rama [1999] for discussions of this literature.) However, as I note later, smaller firms are frequently associated with restatements and fraud, both of which could be associated with internal control problems.

[4] Ernst & Young (1992) identifies the review of internal control as one of the responsibilities of the audit committee, and suggests the committee should engage the independent auditors, the internal auditors, and financial and accounting personnel in these reviews. Similar recommendations under "best practices" for audit committees are made by other Big 5 firms (e.g., PricewaterhouseCoopers 2000; KPMG 1999).

H1: Other things equal, the quality of the audit committee is positively associated with the quality of internal controls.

The mechanisms through which a high-quality audit committee would ensure a high-quality internal control could differ across entities. It could directly engage in overseeing the entity's controls by reviewing "internal accounting procedures and controls with the company's financial and accounting staff," as recommended by a special committee of the AICPA in 1979 (Rittenberg and Nair 1993, 65). More likely, it could engage in discussions with the internal and external auditors about their assessment of internal control. The external auditor is expected (per SAS No. 61, AICPA 1988c) to communicate with the audit committee if it identifies concerns about the quality of the entity's internal controls. A better-quality audit committee is likely to follow up on these communications by eliciting recommendations for improvement and ensuring management carries out these recommendations. Indirect evidence that better audit committees might demand more audit effort or greater audit scope and coverage is provided by Abbott et al. (2003), who report a positive association between audit committee effectiveness and audit fees.

Internal Control Measure

To test the hypothesis, I estimate a logistic regression model with a measure of internal control quality as the dependent variable, and audit committee measures and control variables as independent variables. The internal control quality measure, *ICPROB*, coded 1 or 0 according to the presence or absence of internal control problems, is based on auditor communications that were made privately to the client during the audit, but subsequently made public on the Form 8-K when a change of auditor occurred. Standards for auditor judgment about the entity's internal control are contained in SAS No. 55 (AICPA 1988a), subsequently amended by SAS No. 78 (AICPA 1995). Standards for *communicating* the findings to the client are set out in SAS No. 60 (AICPA 1988b). SAS No. 60 requires that if, during the course of an audit, the auditor becomes aware of deficiencies in an entity's internal control, termed "reportable conditions," it should communicate such matters—through written or oral reports—to the audit committee or to someone of equivalent authority. These deficiencies may be discovered at any point during the audit and must be communicated regardless of the control risk assessment and the decision to rely or not rely on the client's controls.

These communications are not generally publicly available. However, companies that change auditors are required to state whether any "reportable events" had occurred over the previous two years, and if so, disclose the nature of the reportable events. Reportable events include any internal control issues raised by the predecessor auditor (SEC 1988).[5]

In addition to reporting the existence of an internal control problem, the 8-K disclosures also reveal its severity, which, based on SAS No. 60, can take two levels. The first level of problems are "reportable conditions," defined as "significant deficiencies in the design or operation of internal control, which could adversely affect the organization's ability to initiate, record, process, and report financial data consistent with the assertions of management in the financial statements" (SAS No. 60, AICPA 1988b, AU 325). When these reportable conditions are very severe, they are designated "material weaknesses." A material weakness is "a reportable condition in which the design or operation of one or more of the internal control components does not reduce to a relatively low level the risk that misstatements caused by error or fraud in amounts that would be material in relation to the financial statements being audited may occur and not be detected within a timely period by employees in the normal course of performing their assigned functions" (SAS No. 60, AICPA 1988b, AU 325). Thus, a material weakness is a reportable condition, but a reportable condition may not be a material weakness.

---

[5] Reportable events are not the same as reportable conditions. The latter refers specifically to internal control, and the former includes the latter. Financial Reporting Release No. 31 defines reportable events as situations "where the accountant has advised the registrant that it questions the accuracy or reliability of the registrant's financial statements, management's representations, the registrant's internal controls or prior audits" (SEC 1988, 1141). Although the other reportable events may be correlated with the occurrence of internal control problems, they are distinct from the latter.

Audit Committee Quality

Definitions of an effective audit committee vary (DeZoort et al. 2002). One expects both audit committee composition and audit committee activity to be relevant. In general, previous empirical research (Carcello and Neal 2000; Klein 2002b), has focused on composition, possibly because composition is visible while activity is not. In 1999, the major stock exchanges introduced (with endorsement by the SEC) new requirements for audit committee composition and structure (e.g., SEC 1999). All listed companies must appoint audit committees with the following characteristics: (1) at least three members, (2) all members (except in restricted circumstances) independent of management, and (3) at least one member with financial expertise.

Consequently, I use three measures, size (*ACSIZE*), independence (*ACINDEP*), and financial expertise (*ACEXP*) to proxy for audit committee quality. *ACSIZE* is a dummy variable coded 1 if the number of members is three or more, and 0 otherwise. *ACINDEP* is the proportion of members that are independent directors. *ACEXP* is the number of audit committee members with financial expertise. The source for these variables, and their construction, are discussed in Section III.

Other Monitors

Other monitors—management, the external and internal auditors, and the board—interact with the audit committee to determine the entity's internal control. Because the other monitors can also have independent effects on internal control, and one cannot observe the various interactions between these monitors and the audit committee, I include controls representing the other monitors.

Three variables proxy for management quality. First, the extent of formal training of top management in the accounting and finance functions would influence the likelihood of emergence of internal control problems. *ACCTEXP* indicates the presence of personnel, Chief Financial Officer (CFO) or Controller or Chief Accounting Officer (CAO), with a CPA certification or previous experience in a CPA firm. Second, top managers may lack formal training but possess specific experience as overseers of the accounting or finance functions. *CFOEXP* indicates whether the CFO (or Controller or CAO) has had prior experience in one or more of these positions in other employment. I expect both of these variables to be negatively associated with the presence of internal control problems. Third, management integrity is a major determinant of internal control quality. While management attitudes are not observable, I use a proxy, *FRAUD*, based on observed data to capture management integrity. *FRAUD* indicates whether the firm faced SEC action relating to fraudulent financial reporting, and/or restated its financial statements due to GAAP violations, at the time of the 8-K filing or later.[6]

Three auditor characteristics are included, auditor industry share (*ASHARE*), auditor tenure (*TENURE*), and auditor type (*BIG5*). External auditor control variables are needed because, as noted, auditors may have an independent effect on the quality of control due to their charge to assess it. These variables also serve to control for differences in auditors' discovery and reporting of the problem.

Auditor industry specialization (*ASHARE*, measured as the auditor's share in total assets in the firm's two-digit industry) can proxy both for the ability to detect problems due to knowledge of industry-specific issues (Craswell et al. 1995), and its independence.

Prior literature suggests auditor tenure can affect audit quality, auditor independence, and auditor reporting. As auditor tenure increases, client-specific knowledge is enhanced possibly improving audit quality. Myers, Myers, and Omer (2003), Frankel et al. (2002), and Johnson et al. (2002) report negative associations between auditor tenure and measures of discretionary accruals, and Geiger and Raghunandan (2002) report a negative association between auditor tenure and the likelihood of failures in auditor reporting. Therefore, I include auditor tenure (*TENURE*) as a control variable, and predict a negative sign for it.

The Big 5 auditors are widely characterized as "brand-name" and "high-quality" providers, and as the providers of "differentiated" audits (e.g., Craswell et al. 1995; Francis et al. 1999; Krishnan 2003). *A priori*, this suggests they may do better than non-Big 5 auditors in detecting and reporting internal control issues. However, evidence on Big 5/non-Big 5 differences in error detection does not indicate significant differences (DeFond and Jiambalvo 1991; Petroni and Beasley 1996). In addition, clientele differences between the two auditor types—due to differential screening practices by auditors—can result in differences in clients' inherent propensity to generate internal control problems.[7] Thus, I include a dummy variable (*BIG5*) to control for auditor type, but do not predict a sign for it.

A dummy variable for the presence of an internal audit function (*INTAUD*) is included. Because the board can also serve as monitor, and audit committee characteristics can reflect, or be determined by, board composition (Beasley and Salterio 2001; Klein 2002a), I include a board variable, *BDINDEP*, measured as the proportion of nonaudit committee directors that are independent as an additional control (Carcello and Neal 2003). I predict negative signs for both *INTAUD* and *BDINDEP*.

Other Control Variables

Other variables control for a variety of factors that might impact internal control quality. It is generally expected that larger firms have better internal controls (DeFond and Jiambalvo 1991). However, evidence on the association between firm size and accounting errors (which is expected to be negatively correlated with the quality of control) is mixed, as Kinney and McDaniel (1989) note, with some studies reporting a negative association (e.g., Wright and Wright 1996) and others finding no significant difference (e.g., DeFond and Jiambalvo 1991). Studies relating to financial fraud however have generally found a negative association between firm size and the incidence of fraud (e.g., Beasley et al. 2000). Size is measured by the natural logarithm of total assets (*LASSET*). I do not predict a sign for *LASSET*.

---

[6] Other factors, identified in SAS No. 55 (AICPA 1988a), such as ethical values of management or human resource policies and practices, are difficult to measure using archival data. For example, Cohen and Hanno (2000), examine in an experimental setting the effect of corporate governance and management control philosophy on auditors' planning judgments. Other studies examine the impact, in experimental settings, of internal control on audit planning judgments and financial statement errors. See Bell and Wright (1995) for a review of this literature.

[7] Other factors affecting Big 5/non-Big 5 audit quality differences are differences in the provision of internal audit and nonaudit services in conjunction with audit services. Due to lack of data, I do not investigate the effect of these factors.

Firms that grow too rapidly may outgrow their control systems (Kinney and McDaniel 1989; Stice 1991). However, slow growth can also indicate bad performance and a consequent inability to invest in proper controls. Kinney and McDaniel (1989) find, for example, that the incidence of error correction (which is expected to be positively associated with weak internal control) is higher for low-growth companies, possibly indicating bad controls for low-performing (slow-growth) companies rather than high-performing (high-growth) companies. To capture extreme growth levels, high growth (*HIGHGR*) is measured by a dummy variable, coded 1 for growth rate in assets (in the year before auditor change) greater than the industry (two-digit SIC) median, and 0 otherwise. I do not predict a sign for *HIGHGR*. The number of years that a company has been publicly traded has been shown to be associated with fraud and restatements (Beneish 1999; Myers, Myers, Palmrose, and Scholz 2003), which in turn are likely to be associated with weak internal control. I include *IPO*, a variable that indicates whether the company went public within the last five years before the year in which it appears in my sample. Studies also indicate that fraud is most often related to revenue recognition (e.g., Beasley et al. 2000) and restatements are most often related to revenue and receivables (Palmrose and Scholz 2004). Consequently, I control for the ratio of receivables to assets (*RECASST*).

Financial distress may prevent an entity from investing in the maintenance of proper internal control. Distress is measured by two variables: the probability of bankruptcy using the Zmijewski (1984) model (*PRBANK*) and a dummy variable (*LOSS*) indicating a loss during the current year. I include a control for the time period 1999–2000 (*YEAR*), because a number of regulatory changes occurred during this period, some of which could affect audit committee composition. I also include two dummy variables (*NASD* and *NYAM*) indicating the National Association of Securities Dealers Automated Quotation System (NASDAQ)-National Market System and NewYork/American stock exchanges.

Finally, I include a control, *RESIGN* coded 1 for an auditor resignation, and 0 otherwise. The SEC requires disclosure about the nature of auditor change because it believes these disclosures could reveal troublesome audit situations. Previous work (Krishnan and Krishnan 1997; Raghunandan and Rama 1999) has shown that companies whose auditors resign are different from companies whose auditors are dismissed. The inclusion of *RESIGN* controls for these differences.

Choice of Control Sample

The test sample consists of companies that switched auditors and reported internal control problems on their Forms 8-K. The control sample is a sample of auditor switchers that did not report internal control problems on their Forms 8-K.[8] However, one must consider the likelihood that such a sample may in fact include situations where the company has (unreported) internal control problems. The control sample may not be free of internal control problems if (1) the predecessor auditor failed to identify or report internal control problems during its tenure with the client and/or (2) the client failed to report on the Form 8-K that the auditor had, by written or oral report, notified the client of the problem.

---

[8] An alternative control sample is a random sample of companies that did not change auditors. However, since the internal communications about internal control are not publicly available, one cannot ascertain the absence of internal control problems for such a sample.

Failure by the auditor to *identify* an existing internal control problem relating to financial reporting would likely reflect the quality of the audit. Likewise, failure to *report* a problem when identified would depend on auditor quality and independence, and on the auditor's concern with potential litigation and loss of reputation. Controls for auditor differences are included to address this possibility.

As for (2) above, the client is unlikely to avoid reporting an internal control problem on its Form 8-K because the reporting process for auditor changes requires the auditor to endorse the contents of the Form 8-K filed by the client. Further, the reporting requirements state that, in the absence of reportable events, the registrant must specifically state there were no reportable events. This puts pressure on the registrant to make the correct disclosures. Krishnan (2002) finds that situations where the client makes inadequate or inaccurate disclosures (as revealed later by the auditor) on the Form 8-K are relatively infrequent, and that such situations are met with negative stock market reactions. Because clients are likely to be concerned about such negative consequences, and auditors are likely to be concerned about litigation and reputation loss, situations where the problem is not revealed by the client or subsequently by the auditor will be relatively few.

In sum, while one cannot rule out the possibility that the control sample is not free of internal control problems, I expect that such occurrences would be relatively few. Also, any resulting bias would work against finding differences between the test and control samples.[9]

III. DATA

I identified (from the Disclosure Inc. database) an initial sample of companies that changed auditors over the period 1994–2000 and whose Forms 8-K were available on the SEC's EDGAR database.[10] I searched the Forms 8-K for disclosures of internal control problems. This yielded 204 observations. I dropped 40 companies that did not have audit committees. A total of 36 observations were lost because of the unavailability of filings and/or incomplete audit committee or board data (29 observations), or because of incomplete data for one or more of the other variables (seven observations). This yielded an "internal control" (INTCONT) sample of 128 companies.

The financial variables (obtained from the Disclosure database) pertain to the last fiscal year of the predecessor auditor. I constructed the audit committee variables using information in proxy statements, and two of the management quality variables using information in proxy statements and 10-K filings.[11] Following prior work, I define a nonindependent or affiliated director as "current or former officers or employees of the company or of a related entity, relatives of management, professional advisors to the company, (e.g., consultants, bank officers, legal counsel), officers of significant suppliers or customers of the company, and interlocking directors" (Carcello and Neal 2000, 453–454). Using information in the proxies about the employment history of each director, family relationships, and other significant business relationships, I classify audit committee members as independent directors if they are not affiliated (Beasley 1996; Klein 2002b). As in Beasley et al. (1999), I classified each audit committee member as a financial/accounting expert if s/he was a Certified Public Accountant, Chartered Financial Analyst, or had work experience in a financial or accounting position. Using information about past experience of officers, I coded *ACCTEXP* as 1 if the CFO or CAO or Controller had public accounting experience, and 0 otherwise. If the CFO or CAO or Controller had prior experience in a similar capacity, then I coded *CFOEXP* as 1, and 0 otherwise.[12]

---

[9] If the likelihood that the disclosure is contained in the 8-K is systematically associated with audit committee characteristics (for example, because a good audit committee would ensure that the disclosure is made), then audit committee quality could actually be *positively* (rather than *negatively* as hypothesized in this paper) associated with the disclosure of internal control problems. As discussed, it seems likely that the failure to disclose the problem would be related more to auditor quality. Thus, controls for auditor differences control for such potential disclosure differences.

[10] I used the June and December editions for the years 1996–2000 and the June 2001 edition of the Disclosure database. To identify potential auditor changes, I used two techniques: a keyword search, using terms like "resign," "accountant," "auditor change," "dismiss"; and examination of items listed under 8-K item 4 filings on the Disclosure database.

[11] To ensure that the measurement period of the audit committee variables coincided with the period of auditor's assessment of internal control, I examined proxy statements dated close to the auditor change date. On average, the proxy filing dates were 46 (39) days after the date of auditor change for the INTCONT (NOPROB) sample. The median difference between the proxy date and the auditor change date is 53 days for both samples.

To construct *FRAUD*, I searched the SEC website and newswires and newspaper articles using the Lexis/Nexis database. From the SEC website, I obtained information about any action taken by the SEC against the company or its current or former management involving fraud or restatements due to GAAP violations starting from the period of filing of 8-K through January 2004. I searched major newspapers and newswires using keywords such as "restate," "restatement," and "fraud" to identify companies that engaged in fraud or restated their financials due to GAAP violations.

The internal audit variable was constructed, following Beasley et al. (2000), by examining whether the proxies mentioned the presence of an internal audit function.[13] Auditor market share (based on client assets) was constructed for each sample year and for each auditor-two digit SIC industry combination using data from the Disclosure database, using all observations with valid data. I computed the number of years the auditor had been with the auditor (*TENURE*), using Compustat and the Disclosure database for most firms, and proxy filings for a few.[14]

A control sample of companies without internal control problems (NOPROB) was generated from the population of other auditor changes. For each company in the INTCONT sample, a random choice for the control sample was made from among the companies that changed auditors using three criteria: (1) Each INTCONT company was matched with a NOPROB company by industry. Matching by industry is desirable because prior research (Beasley et al. 2000) finds that the likelihood of fraud is higher in certain industries, and there is a direct relation between weak internal control and the incidence of fraud (Bell and Carcello 2000). Industry matching was based on three-digit SIC codes. Where it was not possible to match on three digits, two digits (29 observations) and one digit (14 observations) were used; (2) Each INTCONT company was matched with a NOPROB company on the exchange on which the company was listed. Matching on exchange is necessary because listing requirements relating to the board of directors are different for the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and Over-the-Counter

---

[12] I used keyword searches in proxies and 10-K filings for "chief f," "chief a," or "controller" and variants (e.g., "CFO," "CAO," and "finance") to identify the background information on these individuals. In general, the experience was reported for the past five years, as required by SEC Regulation S-K. For many companies, information was provided for more than five years. Inter-company variation in the number of years covered by these disclosures and the extent of detail about past experience introduces noise in the management quality variables, but such noise is likely to bias the coefficient estimates toward the null hypothesis.

[13] This variable may not completely capture the presence of an internal audit function because its construction relies on the proxies containing nonmandated disclosures about the internal audit function.

[14] For nine companies, I had to rely on the Disclosure database, but because my access to that database goes back only to 1988, the measure of tenure is truncated. However, the results remained qualitatively unchanged when I dropped these nine observations.

(OTC) markets; (3) The control sample was also chosen such that the company did not report any reportable events or disagreements. The final sample has 128 INTCONT companies and 128 NOPROB companies.

IV. EMPIRICAL RESULTS

Table 1, Panels A–C show the sample composition, by year, by industry, and by exchange. The sample ranges from 1994 through 2000, with most observations occurring in 1997 through 1999. The number of observations prior to 1997 is low because this was

TABLE 1
Sample Composition

Panel A: Composition of Sample, By Year of Auditor Change

Year of Change[a]	INTCONT Sample[b]	NOPROB Sample[b]
1994	6	2
1995	14	10
1996	13	21
1997	28	27
1998	20	31
1999	31	27
2000	16	10
Total	128	128

Panel B: Composition of Sample, By Industry

Industry (SIC)	INTCONT[b] Sample	NOPROB[b] Sample
Mining and Construction (10–17)	3	3
Manufacturing (20–39)	39	40
Transportation and Utilities (40–49)	17	16
Wholesale and Retail (50–59)	11	12
Financial Services (60–69)	13	13
Services (70–89)	45	44
Total	128	128

Panel C: Composition of Sample, By Exchange

Exchange	INTCONT[b] Sample	NOPROB[b] Sample
NYSE	16	10
AMEX	8	9
OTC	104	109
Total	128	128

[a] Last fiscal year end of the predecessor auditor.

[b] INTCONT sample consists of companies whose Forms 8-K reporting auditor changes contained disclosures relating to internal control problems. NOPROB sample is a control sample of companies that changed auditors but reported no internal control problems (or any other reportable event or accounting disagreements) in their Forms 8-K.

the phase-in period during which the SEC gradually inducted companies into EDGAR, starting with larger companies. The phase-in period ended in May 1996, after which all companies were required to make filings through EDGAR.[15]

There are some industry-specific patterns in the incidence of internal control problems. Services has the greatest proportion (35.2 percent) followed by Manufacturing (30.5 percent). Within Services (not reported in the table), the incidence of internal control problems is concentrated (20.3 percent) in one two-digit SIC industry (73), Business Services. Within Manufacturing, the incidence of internal control problems is concentrated (19.5 percent) in three two-digit SIC groups, 35 (Industrial Machinery and Equipment), 36 (Electronic and Other Electric Equipment), and 38 (Instruments and Related Products). Panel C of Table 1 indicates that the sample companies are predominantly over-the-counter (OTC) companies: there are 8 (9) companies from the American Stock Exchange (AMEX) in the INTCONT (NOPROB) group, and 16 (10) companies from the New York Stock Exchange (NYSE) in the INTCONT (NOPROB) group.[16]

Despite the attempts to match on industry and exchange, Table 1 indicates that the matching is less than perfect. Two companies (one in Transportation and one in Services) have matches in different industries. However, dropping these two observations and their corresponding matched observations does not change the results I report later. The imperfection in matching is greater with respect to exchange affiliation. In particular, there were not a sufficient number of observations in some industry groups from the NYSE. Therefore, one observation from AMEX and five observations from OTC were chosen as matches for six NYSE observations. I include two dummy variables for exchange affiliation to control for the potential effect of these mismatches.

Table 2 reports descriptive statistics. The univariate tests indicate that the two samples differ in audit committee characteristics. NOPROB companies are more likely to have audit committees of three or more members (*ACSIZE*), have a greater proportion of independent directors (*ACINDEP*) on the audit committee, and a larger number of members with financial expertise (*ACEXP*) than INTCONT companies. NOPROB companies also have a larger proportion (not reported in the tables) of *wholly* independent audit committees (49.2 percent) compared to INTCONT companies (39.8 percent). About 45 percent (61 percent) of the INTCONT (NOPROB) sample (not tabulated) had at least one member with financial expertise. This difference is significant with $p\text{-value} < 0.01$.

Two of the management quality variables, *CFOEXP* and *FRAUD*, are significantly different: INTCONT companies have a lower proportion of CFOs with previous experience in a similar capacity (*CFOEXP*), and are more likely to be associated with SEC action or restatements (*FRAUD*). INTCONT companies are financially more stressed, as measured by *LOSS*, have shorter auditor tenures (*TENURE*), and are more likely to have had IPOs.

---

[15] The industry and exchange matching procedure minimizes any selectivity bias that may arise from the inclusion of companies from the phase-in period. Also, models that were estimated after eliminating the observations from the phase-in period yielded results that were similar to the reported results.

[16] Further information for the OTC companies was not always available. Where information could be obtained, 58 (50) of the INTCONT (NOPROB) companies were traded on the NASDAQ-National Market System, 16 (24) on the NASDAQ-Small-Cap Market, and 18 (16) on the Bulletin Board and Pink Sheets. For the remaining companies the information available from Forms 10-K or newspaper articles indicated only that it was traded either in NASDAQ (12 and 15 companies in the INTCONT and NOPROB samples, respectively) or in the Over-the-Counter market (4 companies in the NOPROB sample).

TABLE 2
Descriptive Statistics for INTCONT and NOPROB Samples

Variable	Expected Sign	INTCONT[a] Sample		NOPROB[a] Sample		t-statistic[b]	Wilcoxon Z[c]
		Mean	Median	Mean	Median		
ACSIZE	−	0.547	1	0.641	1	−1.53*	−1.52*
ACINDEP	−	0.681	0.670	0.771	0.815	−2.54***	−2.24**
ACEXP	−	0.531	0	0.789	1	−2.91***	−2.72***
ACCTEXP	−	0.219	0	0.234	0	−0.30	−0.30
CFOEXP	−	0.508	1	0.680	1	−2.83***	−2.79***
FRAUD	+	0.242	0	0.055	0	4.36***	4.21***
ASHARE	+	0.145	0.100	0.116	0.072	1.26	1.94**
TENURE	−	5.032	4	6.492	4	−2.07**	−1.80**
BIG5	+/−	0.859	1	0.773	1	1.78*	1.77*
INTAUD	−	0.188	0	0.211	0	−0.47	−0.47
BDINDEP	−	0.323	0.333	0.347	0.333	−0.70	−0.83
LASSET	+/−	4.148	3.962	3.787	3.600	1.55	1.52
HIGHGR	+/−	0.516	1	0.492	0	0.37	0.37
IPO	+	0.430	0	0.352	0	1.28*	1.28*
RECASST	+	0.230	0.209	0.226	0.181	0.19	0.89
PRBANK	+	0.105	0.0002	0.099	0.00003	0.19	1.74**
LOSS	+	0.727	1	0.586	1	2.39***	2.36***
YEAR	+/−	0.367	0	0.289	0	1.33	1.33
NASD	+/−	0.453	0	0.391	0	1.01	1.01
NYAM	+/−	0.188	0	0.148	0	0.83	0.83
RESIGN	+	0.445	0	0.117	0	6.25***	5.83***
n		128		128			

***, **, * Significant at 1 percent, 5 percent, and 10 percent levels, respectively, one-tailed where signs are predicted, two-tailed otherwise.

[a] INTCONT sample includes companies whose Forms 8-K reporting auditor changes contained disclosures relating to internal control problems. NOPROB sample is a control sample of companies that changed auditors but reported no internal control problems (or any other reportable event or accounting disagreements) in their Forms 8-K.

[b] Tests for significant differences in means.

[c] Tests whether the observations in the two groups are from populations with different distribution.

Variable Definitions:

ACSIZE = 1 if audit committee has at least three members, and 0 otherwise;
ACINDEP = proportion of independent outside members on the audit committee;
ACEXP = number of audit committee members with financial expertise;
ACCTEXP = 1 if the Chief Financial Officer (or Chief Accounting Officer or Controller) has CPA certification or previous experience in public accounting, and 0 otherwise;
CFOEXP = 1 if the Chief Financial Officer (or Chief Accounting Officer or Controller) has previous experience in a similar capacity with another company, and 0 otherwise;
FRAUD = 1 if there was an SEC action against the company's management or the company restated financials due to GAAP violation, and 0 otherwise;
ASHARE = auditor's share in two-digit industry;
TENURE = number of years the auditor has audited the client;
BIG5 = 1 if the predecessor auditor is Big 5, and 0 otherwise;
INTAUD = 1 if internal audit function exists, and 0 otherwise;
BDINDEP = proportion of independent outside members on the board, excluding the audit committee;
LASSET = natural logarithm of total assets (in million);

(continued on next page)

TABLE 2 (continued)

HIGHGR = 1 if growth rate in assets above the industry median, and 0 otherwise;
IPO = 1 if the company had an *IPO* within five years of the last year of the predecessor auditor, and 0 otherwise;
RECASST = ratio of receivables to total assets;
PRBANK = probability of bankruptcy using Zmijewski's (1984) model;
LOSS = 1 if net income is negative, and 0 otherwise;
YEAR = 1 if the observation is from 1999 or 2000, and 0 otherwise;
NASD = 1 if the stock was traded in National Market System of NASDAQ, and 0 otherwise;
NYAM = 1 if the stock was traded in NYSE or AMEX, and 0 otherwise; and
RESIGN = 1 if the auditor resigned, and 0 otherwise.

INTCONT companies are more likely to have Big 5 auditors than NOPROB companies.[17] Auditor resignations are also more likely for INTCONT companies.[18] Finally, the INTCONT sample has higher medians but not means for *ASHARE*, and *PRBANK*.

Table 3 presents correlations. All correlations between the independent variables are below +/− 0.4. Further, investigation of variance inflation factors—all of which were under 1.6—suggests multicollinearity is not a concern.

Table 4 presents logistic regression estimates for the incidence of internal control problems. The dependent variable, *ICPROB* is coded 1 or 0 according to the presence or absence of internal control problems. The model is significant, as indicated by its Chi-squared value. Two kinds of tests are used to examine the significance of coefficient estimates. Column 2 presents p-values based on conventional logistic estimates. Column 3 presents the results of nonparametric randomization tests, which examine the generic hypothesis that a set of variables is unrelated to the dependent variable (Noreen 1989). The test is conducted by shuffling the dependent variable randomly across the observations and estimating the logit model. This process is repeated 1,000 times, generating a distribution of coefficients that can be compared with the original coefficients estimated with the unshuffled data. If the coefficient is predicted to be positive, then the significance of the coefficient (randomized p-value) is given by $(NGE + 1)/(NS + 1)$ where *NGE* is the number of (shuffled) coefficients that exceed the original coefficient, and *NS* is the number of times the shuffling is done. If the coefficient is predicted to be negative, *NGE* is defined as the number of times the (shuffled) coefficients is less than the original coefficient (DeFond and Jiambalvo 1991). The test statistic is not constructed for coefficients for which a sign is not predicted.

A comparison of columns 2 and 3 in Table 4 indicates that the results of the randomization tests closely parallel the conventional tests. The proportion of independent members on the audit committee (*ACINDEP*) and the number of expert members (*ACEXP*) are negatively associated with the likelihood of an internal control problem (p-value = 0.005 and 0.014, respectively, in column 2, and p-values of 0.002 and 0.011, respectively, in column 3). Two of the management quality proxies (*CFOEXP* and *FRAUD*) are strongly significant. Thus situations where the CFO (or equivalent officer) has had prior experience in a similar

---

[17] If the Big 5 report internal control problems more often, then the difference in the means suggests that the NOPROB sample may have undetected or unreported internal control problems. This shortcoming in the data cannot be overcome, but would bias against my finding significant differences between the samples. Also, as I report later, I estimate the models for the Big 5 alone, and find similar results.

[18] I examined univariate differences among INTCONT and NOPROB companies within the resignation and dismissal subsamples. There are two differences. The proportion of companies with three or more audit committee members (*ACSIZE*) is significantly higher for dismissals as compared to resignations for both the INTCONT and NOPROB companies. A greater proportion of INTCONT resignation companies have high growth (*HIGHGR*) as compared to INTCONT dismissal companies.

TABLE 3
Correlations

	ACSIZE	*ACINDEP*	*ACEXP*	*ACCTEXP*	*CFOEXP*	*FRAUD*	*ASHARE*	*TENURE*	*BIG5*	*INTAUD*	*BDINDEP*
ACSIZE	1.00	−0.08	0.16***	−0.01	−0.02	−0.08	−0.05	0.10	−0.13**	0.09	0.04
ACINDEP		1.00	−0.02	−0.002	−0.04	−0.06	−0.03	0.10	−0.01	0.07	−0.06
ACEXP			1.00	0.07	0.04	0.01	−0.11*	0.05	−0.10	0.06	0.02
ACCTEXP				1.00	0.16***	0.01	−0.03	−0.04	0.04	−0.04	−0.12*
CFOEXP					1.00	0.01	−0.08	−0.12*	−0.02	−0.01	0.06
FRAUD						1.00	0.03	−0.03	0.11*	−0.07	0.01
ASHARE							1.00	−0.02	0.34***	0.04	0.09
TENURE								1.00	0.13**	0.09	0.01
BIG5									1.00	0.06	0.20***
INTAUD										1.00	0.05
BDINDEP											1.00

(*continued on next page*)

TABLE 3 (continued)

	LASSET	*HIGHGR*	*IPO*	*RECASST*	*PRBANK*	*LOSS*	*YEAR*	*NASD*	*NYAM*	*RESIGN*
ACSIZE	0.12**	−0.03	−0.12*	0.11*	−0.05	−0.10	0.002	0.01	0.05	−0.21***
ACINDEP	−0.01	−0.05	−0.05	0.04	0.07	−0.003	0.21***	0.04	−0.03	−0.05
ACEXP	−0.02	−0.002	0.05	0.05	−0.02	−0.09	−0.09	0.05	−0.02	−0.13**
ACCTEXP	−0.02	0.15**	0.01	−0.04	−0.004	0.04	−0.04	0.01	0.01	−0.03
CFOEXP	0.05	0.05	0.19***	−0.09	−0.05	0.04	−0.02	−0.03	0.01	−0.10
FRAUD	0.07	0.04	0.05	0.01	−0.12*	0.05	0.01	0.02	0.05	0.13**
ASHARE	0.03	0.08	0.02	0.07	0.08	0.04	0.07	−0.02	−0.03	0.02
TENURE	0.26***	−0.10	−0.39***	−0.003	−0.10	−0.16**	−0.002	−0.001	0.17***	−0.15**
BIG5	0.31***	0.03	0.03	−0.04	−0.09	0.08	−0.03	0.14**	0.13**	0.03
INTAUD	0.15**	0.05	0.002	0.11*	0.04	−0.24***	−0.06	−0.05	0.09	0.01
BDINDEP	0.18***	−0.04	0.07	0.01	−0.02	0.04	0.09	0.09	0.01	−0.10
LASSET	1.00	0.03	−0.01	0.17***	−0.37***	−0.27***	−0.01	0.19***	0.33***	−0.01
HIGHGR		1.00	0.12*	−0.04	−0.01	−0.01	−0.01	−0.09	0.05	0.12*
IPO			1.00	0.06	−0.02	0.06	−0.12*	0.18***	−0.15**	0.14**
RECASST				1.00	−0.15**	−0.29***	−0.12*	0.05	−0.13**	0.04
PRBANK					1.00	0.27***	0.16***	- 0.14**	−0.04	−0.01
LOSS						1.00	0.19***	−0.09	−0.01	−0.01
YEAR							1.00	−0.04	−0.07	0.14**
NASD								1.00	−0.38***	0.03
NYAM									1.00	−0.05
RESIGN										1.00

***, **, * Significant at 1 percent, 5 percent, and 10 percent levels, respectively, two-tailed tests.
See Table 2 for definition of variables.

TABLE 4
Logistic Regression for the Incidence of Internal Control Problems[a]

Variable	Expected Sign	Coefficient Estimate (1)	Logistic p-value (2)	Randomization Test p-value (3)
Intercept	+/−	−0.074	0.927	—
ACSIZE	−	−0.095	0.387	0.378
ACINDEP	−	−1.467***	0.005	0.002
ACEXP	−	−0.487**	0.014	0.011
ACCTEXP	−	0.091	0.404	0.599
CFOEXP	−	−0.917***	0.003	0.005
FRAUD	+	1.769***	0.000	0.000
ASHARE	+	0.363	0.351	0.351
TENURE	−	−0.055*	0.082	0.030
BIG5	+/−	0.036	0.939	—
INTAUD	−	0.199	0.313	0.713
BDINDEP	−	−0.496	0.191	0.167
LASSET	+/−	0.230**	0.045	—
HIGHGR	+/−	−0.207	0.513	—
IPO	+	0.076	0.418	0.401
RECASST	+	0.169	0.424	0.416
PRBANK	+	0.447	0.233	0.190
LOSS	+	0.875***	0.009	0.006
YEAR	+/−	0.207	0.568	—
NASD	+/−	0.341	0.351	—
NYAM	+/−	0.297	0.562	—
RESIGN	+	1.668***	0.000	0.000
Chi-Squared[b] (p-value)			88.567 (0.000)	
Pseudo R^2			0.250	
n			256	

***, **, * Significant at 1 percent, 5 percent, and 10 percent levels, respectively, one-tailed where signs are predicted, two-tailed otherwise.

[a] Dependent Variable: *ICPROB* = 1 for INTCONT companies, and 0 for NOPROB companies.

[b] The Chi-squared statistic and the Pseudo R^2 are for the logistic regression model. The coefficient estimates and the corresponding p-values for the logistic regression model are reported in columns 1 and 2. The p-values reported in column 3 are generated using Noreen's (1989) randomization procedure. See text for details.

See Table 2 for definition of independent variables.

capacity are less likely to be associated with internal control problems. Situations of potential fraud are characterized by a higher incidence of internal control problems.[19]

---

[19] To test whether the effect of the audit committee variables varied with management quality and/or resignation status, I also examined the effects of interactions between (1) the audit committee variables (*ACEXP* and *ACINDEP*) and the management quality variables (*CFOEXP* and *FRAUD*), and (2) the audit committee variables (*ACEXP* and *ACINDEP*) and *RESIGN*. None of the interactions was significant. The audit committee variables remained significant. Further, I examined the effect of the interactions between (3) the two management quality variables, *CFOEXP* and *FRAUD*, and (4) the management quality variables (*CFOEXP* and *FRAUD*) and *RESIGN*. Again, none of the interactions was significant. The results pertaining to audit committee variables remained unchanged.

Among the other control variables, financial distress (measured by *LOSS*) is positively associated with, and auditor tenure is negatively associated with, the likelihood of an internal control problem. Consistent with the univariate results, firm size (*LASSET*) is positively associated with the likelihood of internal control problems. This runs counter to the generally expressed view that larger firms have better controls, although as noted earlier there is no consistent evidence to support this view. It is possible that the expected association between control quality and firm size does not hold for small and medium sized companies that comprise the current sample. Finally, as indicated in the univariate analysis, resignations (*RESIGN*) are positively associated with internal control problems.

Severity of Internal Control Problem

Next, I examine the severity of the internal control problems, that is, whether the problem is a reportable condition (REPCOND) or a material weakness (MATWEAK). Of the 128 companies in the INTCONT sample, 74 had material weaknesses relating to internal control. A multinomial logit model is presented in Table 5. The dependent variable, *ICSEV*, is coded 2 for a material weakness, 1 for a reportable condition, and 0 for no problems. *A priori*, one would expect a negative association between variables representing audit committee effectiveness and the severity of the internal control problem. Columns 1–6 present the coefficients for the effects of the variables on the likelihood of a reportable condition (columns 1–3) and material weakness (columns 4–6) *relative* to the likelihood of there not being a problem. In column 7, I present the difference in the coefficients to test whether the coefficient is significantly different across the two levels. Because the power of these tests is generally lower than that of the logit (due to a smaller number of observations in subcategories of the data), I focus only on the results that are significant according to both the conventional t-test and the randomization tests.

Results in columns 1–6 indicate that, with the exception of *TENURE*, which is significant for reportable conditions but not for material weaknesses, the results in Table 4 carry over to both levels of internal control problems. *ACINDEP* and *ACEXP* are significant. *CFOEXP* and *FRAUD* are negatively and positively associated with both levels of internal control problems, respectively. Firms with losses are more likely to have both reportable conditions and material weaknesses relative to NOPROB companies. *RESIGN* is positively associated with both kinds of internal control problems.[20]

Column 7 indicates that there are relatively few differences between the coefficients for the two kinds of problems. Audit committee characteristics do not affect the relative likelihood of a material weakness versus a reportable condition. Thus, while audit committee independence and expertise lower the likelihood of an internal control problem, they do not affect the relative likelihood of one type of problem compared to the other. Three factors are more likely to be related to material weakness than to reportable conditions: *FRAUD*, *BIG5*, and *PRBANK*. Thus, the severity of the internal control problem is greater, not surprisingly, in situations of subsequent fraud, and greater financial distress. In addition, Big 5 firms are more likely to be associated with material weaknesses than with reportable

---

[20] The definition of financial expertise used in this paper is narrower than the BRC definition, but similar to the SEC's (1999) definition. For the BRC, a financial expert also includes "being or having been a CEO or other senior officer with financial oversight responsibilities" (BRC 1999, 25). As a result, the incidence of financial expertise in my sample is lower than that in studies that use the BRC's definition (e.g., Carcello and Neal 2003). As a sensitivity exercise, I re-estimated the models using the BRC definition. *ACEXP* is negative and significant in the logit, and negative and significant for material weaknesses and insignificant for reportable conditions in the multinomial logit.

TABLE 5
Multinomial Logit[a]

Variable	Expected Sign	REPCOND[b] vs. NOPROB Coefficient Estimate (1)	REPCOND[b] vs. NOPROB Logistic p-value (2)	REPCOND[b] vs. NOPROB Randomization Test p-value (3)	MATWEAK[b] vs. NOPROB Coefficient Estimate (4)	MATWEAK[b] vs. NOPROB Logistic p-value (5)	MATWEAK[b] vs. NOPROB Randomization Test p-value (6)	Difference in Coefficients [Column (4) minus Column (1)][c] (7)
Intercept	+/−	0.224	0.815	—	−1.748*	0.081	—	−1.972
ACSIZE	−	−0.253	0.268	0.460	−0.039	0.459	0.290	0.214
ACINDEP	−	−1.697***	0.007	0.022	−1.221**	0.031	0.007	0.476
ACEXP	−	−0.404*	0.071	0.015	−0.529**	0.022	0.074	−0.125
ACCTEXP	−	0.134	0.387	0.570	0.088	0.419	0.602	−0.046
CFOEXP	−	−1.119***	0.003	0.021	−0.737**	0.026	0.003	0.382
FRAUD	+	1.280**	0.020	0.000	2.070***	0.000	0.023	0.790*
ASHARE	+	1.368	0.104	0.696	−0.473	0.340	0.138	−1.841
TENURE	−	−0.113**	0.030	0.014	−0.024	0.276	0.217	0.089
BIG5	+/−	−0.754	0.169	—	0.958	0.137	—	1.712***
BDINDEP	−	−0.316	0.331	0.114	−0.731	0.131	0.336	−0.415
INTAUD	−	0.634	0.101	0.357	−0.164	0.370	0.895	−0.798
LASSET	+/−	0.144	0.320	—	0.268**	0.041	—	0.124
HIGHGR	+/−	−0.160	0.681	—	−0.268	0.463	—	−0.108
IPO	+	−0.187	0.342	0.281	0.232	0.290	0.665	0.419
RECASSET	+	1.303	0.109	0.845	−0.949	0.198	0.130	−2.252
PRBANK	+	−0.602	0.240	0.056	1.098*	0.053	0.748	1.700**
LOSS	+	0.895**	0.027	0.012	0.865**	0.024	0.044	−0.030
YEAR	+/−	0.194	0.678	—	0.175	0.667	—	−0.019
NASD	+/−	0.305	0.507	—	0.420	0.328	—	0.115
NYAM	+/−	0.482	0.448	—	0.278	0.633	—	−0.204
RESIGN	+	1.769***	0.000	0.000	1.594***	0.000	0.000	−0.175
Chi-squared[d]					118.585			
(p-value)					(0.000)			
Pseudo R^2					0.224			
n					256 (54 REPCOND, 74 MATWEAK, and 128 NOPROB)			

TABLE 5 (continued)

***, **, * Significant at 1 percent, 5 percent, and 10 percent levels, respectively, one-tailed where signs are predicted, two-tailed otherwise.

[a] Dependent Variable: *ICSEV* = 2 for MATWEAK companies, 1 for REPCOND companies, and 0 for NOPROB companies.

[b] MATWEAK are INTCONT companies with material weaknesses. REPCOND are INTCONT companies with reportable conditions. Columns 1–3 model the probability of REPCOND relative to the probability of NOPROB. Columns 4-6 model the probability of MATWEAK relative to the probability of NOPROB.

[c] Models the probability of MATWEAK relative to the probability of REPCOND. Significance is based on the randomization test p-values.

[d] The Chi-squared statistic and the Pseudo R^2 are for the multinomial logit model. The coefficient estimates and the corresponding p-values for the multinomial logit model are reported in columns 1, 2, 4, and 5. The p-values reported in columns 3 and 6 are generated using Noreen's (1989) randomization procedure. See text for details.

See Table 2 for definition of the independent variables.

conditions. This could reflect differences in reporting or differences in clientele across Big 5 and non-Big 5 firms.

V. ADDITIONAL TESTS

Alternative Audit Committee Quality Measures

I examined other measures of audit committee effectiveness. First, I classified the audit committee as being of high quality (*INDEP*) if all members are independent (Klein 2002b). Second, I classified an audit committee as being of high quality (*INDEP-EXP*) if it has all independent members and at least one member with financial expertise (Raghunandan et al. 2001). Logistic regressions (Panel A, columns 1–2) and multinomial logits (Panel B, columns 1–4) using these two alternatives are presented in Table 6. The results for the other variables are similar in all cases and therefore not reported in this table. Both panels indicate that *INDEP* and *INDEP-EXP* are significant and negative.

Auditor Dismissals

Tables 4 and 5 indicated that auditor resignations are positively associated with internal control problems. As a sensitivity test, I also estimated the models using the sample of dismissals alone. I retained only pairs of observations, that is, INTCONT companies for which the matching NOPROB companies were also dismissal companies. The logistic and multinomial logit results are presented in Table 6, Panel A (column 3) and Panel B (columns 5–6). Panel A indicates, consistent with Table 4, both *ACINDEP* and *ACEXP* are negative and significant, suggesting that the type of auditor change may not be driving the full sample results. In Panel B, *ACINDEP* is significant for both reportable conditions and material weaknesses and *ACEXP* is significant for reportable conditions but not for material weaknesses.

Auditor Type

As noted, the sample consists predominantly of Big 5 clients. To check that the results are not driven by the few non-Big 5 clients, I also estimated the models for the sample of Big 5 clients only (once again retaining only matched pairs). The results are reported in Table 6, Panels A (column 4) and B (columns 7–8). The results are generally consistent with Tables 4 and 5. *ACINDEP* is significant and negative in both panels. *ACEXP* is significant in Panel A and significant for material weaknesses in Panel B.

Additional Audit Committee Characteristics

Next, I extend the measures of audit committee effectiveness to audit committee characteristics other than independence and expertise. I add three variables suggested in prior work as additional measures of audit committee effectiveness: number of meetings (Abbott and Parker 2000), proportion of stock owned by audit committee members, and the average number of directorships held by the audit committee members in other companies (Carcello and Neal 2003).

The expected signs for all three variables are ambiguous. The number of meetings may reflect how proactive the audit committee is in assessing internal control, but it may also be a response to the emergence of problems. Stock ownership by directors can have both positive and negative effects on director behavior. The positive effect, derived from agency theory, suggests ownership provides incentives to take actions to protect one's investment, including those that ensure proper financial reporting. The negative effect is the incentive to manipulate stock prices through inaccurate reporting. The number of directorships held

TABLE 6
Sensitivity Analyses

Panel A: Logistic Regression Models[a]

Variable	Expected Sign	Coefficient Estimate (p-value) *INDEP* (1)	 *INDEP-EXP* (2)	 Dismissals Only (3)	 Big 5 Only (4)
Intercept	+/−	−0.776 (0.276)	−1.277* (0.060)	0.695 (0.546)	−0.906 (0.397)
ACSIZE	−	−0.171 (0.309)	−0.108 (0.368)	0.101 (0.415)	0.219 (0.308)
ACINDEP	−	—	—	−1.995*** (0.009)	−1.809** (0.011)
INDEP	−	−0.691** (0.021)	—	—	—
INDEP-EXP	−	—	−0.908*** (0.009)	—	—
ACEXP	−	−0.499** (0.012)	—	−0.748** (0.014)	−0.577** (0.024)
Chi-Squared (p-value)		85.932 (0.000)	82.619 (0.000)	34.536 (0.023)	69.218 (0.000)
Pseudo R^2		0.242	0.233	0.189	0.294
n		256	256	132	170

(*continued on next page*)

TABLE 6 (continued)

Panel B: Multinomial Logit[b]

		Coefficient Estimate (p-value)							
		INDEP		*INDEP-EXP*		Dismissals Only		Big 5 Only	
Variable	**Expected Sign**	**REPCOND (1)**	**MATWEAK (2)**	**REPCOND (3)**	**MATWEAK (4)**	**REPCOND (5)**	**MATWEAK (6)**	**REPCOND (7)**	**MATWEAK (8)**
Intercept	+/−	−0.623	−2.312**	−1.084	−2.820***	1.038	−1.300	−1.940	−1.371
		(0.462)	(0.012)	(0.183)	(0.001)	(0.470)	(0.385)	(0.165)	(0.255)
ACSIZE	−	−0.339	−0.112	−0.258	−0.064	−0.140	0.293	0.005	0.311
		(0.210)	(0.387)	(0.259)	(0.430)	(0.413)	(0.294)	(0.496)	(0.261)
ACINDEP	−	—	—	—	—	−2.594***	−1.805**	−1.664**	−1.924**
						(0.008)	(0.029)	(0.045)	(0.013)
INDEP	−	−0.702**	−0.633*	—	—	—	—	—	—
		(0.047)	(0.052)						
INDEP-EXP	−	—	—	−0.955**	−0.797**	—	—	—	—
				(0.027)	(0.038)				
ACEXP	−	−0.410*	−0.546**	—	—	−1.199***	−0.368	−0.241	−0.807***
		(0.068)	(0.019)			(0.006)	(0.178)	(0.250)	(0.010)
Chi-squared		115.541		112.172		61.223		90.979	
(p-value)		(0.000)		(0.000)		(0.017)		(0.000)	
Pseudo R^2		0.218		0.212		0.226		0.262	
n		256		256		132		170	

***, **, * Significant at 1 percent, 5 percent, and 10 percent levels, respectively, one-tailed where signs are predicted, two-tailed otherwise.

[a] Dependent Variable: *ICPROB* = 1 for INTCONT companies, and 0 for NOPROB companies. Other variables included are the same as in Table 4 except that *RESIGN* is not included in column 3 and *BIG5* is not included in column 4. Only audit committee variable estimates are reported.

[b] Dependent Variable: *ICSEV* = 2 for MATWEAK companies, 1 for REPCOND companies, and 0 for NOPROB companies. Other variables included are the same as in Table 5 except that *RESIGN* is not included in columns 5–6 and *BIG5* is not included in columns 7–8. Only audit committee variable estimates are reported.

INDEP = 1 if all members are independent, 0 otherwise. *INDEP-EXP* =1 if all members are independent and at least one member has financial expertise, 0 otherwise. See Table 2 for definition of other independent variables.

by audit committee members on the boards of other corporations is one element of *governance* expertise. On the one hand, the greater the number of outside directorships, the more the audit committee member's investment in obtaining governance expertise, and therefore more likely s/he will engage in effective monitoring activities. On the other hand, an increasing number of outside directorships can cause attention to be spread out thinly over the different assignments, thus lowering efficacy in any one firm.

The three variables are not significant in the logistic or multinomial logits (results not tabulated), perhaps reflecting the fact that both positive and negative forces are at work. Other results remain unchanged.

VI. DISCUSSION AND CONCLUSIONS

Recent policy initiatives have focused on (1) improving corporate internal control, and (2) emphasizing more formally than before the role of the audit committee in maintaining effective control. An example of the first kind of policy initiative is the requirement that auditors attest to management's assertions about internal control (SEC 2003a).[21] Examples of the second kind of policy are (1) the recent mandate that all auditor communication on internal control be specifically provided to the audit committee (SEC 2003b) and (2) the new SEC definition of an "audit committee expert," which includes familiarity with internal control issues as a component (SEC 2003c). Policy makers and the profession have always maintained that the audit committee should adopt the monitoring of control as a primary function. Yet, because internal control quality is not observable, little is known about how entities achieve high-quality internal control (Kinney 2000), and whether audit committees play a role (Cohen et al. 2002).

In this paper, I provide evidence of (1) a negative association between the presence of internal control problems and audit committee independence and (2) a negative association between the presence of internal control problems and the number of audit committee members with financial expertise.[22] In addition, I find four other factors to be consistently associated with the incidence of internal control problems: managers' (CFO, CAO, or Controller) past on-the-job experience, management's tendency to engage in fraud, auditor tenure, and financial stress. Variables representing other monitors, such as the internal audit function, the external auditor, and the board of directors are not consistently related to the quality of internal control. When I partition the internal control problems into less severe (reportable conditions) and more severe (material weaknesses) problems, I find that the audit committee associations reported above hold for both levels of problems in the basic model. In sensitivity analyses, the negative significance of audit committee independence persists in all cases. Audit committee expertise is significant for at least one of the two levels of internal control problems in all sensitivity exercises.

The results are consistent with the current policy emphasis on the independence and expertise of audit committee members. The adequacy of both requirements has recently come under criticism (Lublin 2002; Whitehead 2002). The results pertaining to expertise are particularly interesting because a substantial proportion (55 percent and 39 percent of internal control and no-problem firms, respectively) of my sample did not have any member with expertise. The 1999 exchange rules (effective June 2001) can be expected to bring changes in audit committee composition as a result of companies complying with the requirement to appoint at least one financial expert to the audit committee. Moreover, the SEC's new definition of audit committee expertise includes as noted, "experience with internal controls and procedures for financial reporting" (SEC 2003c). Future studies can examine whether the enhancement to the expertise definition results in a closer association between audit committee quality and internal control quality.

---

[21] So far, management reports on internal control have been voluntary rather than mandated (Hermanson 2000).

[22] Because the sample time period precedes the effective dates for recent SEC policies regarding audit committee composition, there is enough variation in the data to allow inferences about the association between audit committee composition and internal control quality.

A shortcoming of this research, also noted by previous archival research in this area (e.g., Carcello and Neal 2003), is that it is a study of association, not causation. It is vulnerable to the possibility that unobserved factors that are correlated with both internal control and audit committee characteristics may drive the reported findings. To minimize this possibility, I use control variables representing other monitors of internal control, including the internal audit function, the external auditor, the board of directors, and the quality of management, as well as a number of variables (such as financial stress and growth) that may be related to the quality of internal control and the quality of the audit committee. The finding that the audit committee characteristics are associated with internal control after controlling for other monitors, suggests the audit committee has a contribution over and above the other monitors.

A second shortcoming is that the sample consists of companies that changed auditors. Because reported internal control problems are only made public in situations of auditor change, the sample choice was unavoidable. Auditor change companies tend to be smaller in size than the general population of companies and are traded on the smaller stock exchanges. This may affect the generalizability of the results of this study. However, it is generally small firms and firms on the NASDAQ and OTC exchanges that are vulnerable to fraudulent reporting, earnings restatements, and SEC action (Dechow et al. 1996; Bonner et al. 1998). To the extent that weak internal control *contributes* to these negative outcomes, this study focuses on the potential role of the monitoring provided by audit committees in avoiding such outcomes. My results are consistent with prior work that also uses restricted settings to test the effect of audit committees—for example, financially stressed firms (Carcello and Neal 2000) or auditor change firms (Carcello and Neal 2003)—as well as with Klein (2002b), which is based on a more general population of firms.

Under the new SEC rules (SEC 2003a), auditors will issue attestation reports on their clients' internal control. Because these reports must be filed by all filers and not just those changing auditors, future studies can examine the role of audit committee composition and other factors in determining the quality of internal control for a more general population of firms.

REFERENCES

Abbott, L. J., and S. Parker. 2000. Auditor selection and audit committee characteristics. *Auditing: A Journal of Practice & Theory* 19 (Fall): 47–66.

———, ———, G. F. Peters, and K. Raghunandan. 2003. The association between audit committee characteristics and audit fees. *Auditing: A Journal of Practice & Theory* 22 (September): 17–32.

American Institute of Certified Public Accountants (AICPA). 1988a. *Consideration of the Internal Control Structure in a Financial Statement Audit.* Statement on Auditing Standards No. 55. New York, NY: AICPA.

———. 1988b. *Communication of Internal Control Structure Related Matters Noted in an Audit.* Statement on Auditing Standards No. 60. New York, NY: AICPA.

———. 1988c. *Communication with Audit Committees*. Statement on Auditing Standards No. 61. New York, NY: AICPA.

———. 1995. *Consideration of Internal Control in a Financial Statement Audit: An Amendment to Statement on Auditing Standards No. 55*. Statement on Auditing Standards No. 78. New York, NY: AICPA.

Beneish, M. D. 1999. Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74 (October): 425–457.

Beasley, M. S. 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review* 71 (October): 443–465.

———, J. V. Carcello, and D. R. Hermanson. 1999. *Fraudulent Financial Reporting: 1987–1997, An Analysis of U.S. Public Companies*. New York, NY: COSO.

———, ———, ———, and P. D. Lapides. 2000. Fraudulent financial reporting: Consideration of industry traits and corporate governance mechanisms. *Accounting Horizons* 14 (December): 441–454.

———, and S. E. Salterio. 2001. The relationship between board characteristics and voluntary improvements in audit committee composition and experience. *Contemporary Accounting Research* 18 (Winter): 539–570.

Bell, T. B., and A. M. Wright. 1995. *Auditing Practice, Research and Education: A Productive Collaboration*. New York, NY: AICPA.

———, and J. V. Carcello. 2000. A decision aid for assessing the likelihood of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory* 19 (Spring): 169–184.

Blue Ribbon Committee (BRC). 1999. *Report and Recommendations of the Blue Ribbon Committee on Improving the Effectiveness of Corporate Audit Committees*. New York, NY: NYSE and NASD.

Bonner, S. E., Z-V. Palmrose, and S. M. Young. 1998. Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases. *The Accounting Review* 73 (October): 503–532.

Carcello, J. V., and T. L. Neal. 2000. Audit committee composition and auditor reporting. *The Accounting Review* 75 (October): 453–467.

———, D. R. Hermanson, and T. L. Neal. 2002. Disclosures in audit committee charters and reports. *Accounting Horizons* 16 (December): 291–304.

———, and T. L. Neal. 2003. Audit committee characteristics and auditor dismissals following "new" going-concern reports. *The Accounting Review* 78 (January): 95–117.

Cohen, J. R., and D. M. Hanno. 2000. Auditors' consideration of corporate governance and management control philosophy in preplanning and planning judgments. *Auditing: A Journal of Practice & Theory* 19 (Fall): 133–146.

———, G. Krishnamoorthy, and A. M. Wright. 2002. Corporate governance and the audit process. *Contemporary Accounting Research* 19 (Winter): 573–594.

Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20 (December): 297–322.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (Spring): 1–36.

DeFond, M. L., and J. Jiambalvo. 1991. Incidence and circumstances of accounting errors. *The Accounting Review* 66 (July): 643–655.

DeZoort, F. T. 1997. An investigation of audit committees' oversight responsibilities. *Abacus* 33 (September): 208–227.

———. 1998. An analysis of experience effects on audit committee members' oversight judgments. *Accounting, Organizations and Society* 23 (January): 1–21.

———, D. R. Hermanson, D. S. Archambeault, and S. A. Reed. 2002. Audit committee effectiveness: A synthesis of the empirical audit committee literature. *Journal of Accounting Literature* 21: 38–75.

Ernst & Young. 1992. *Audit Committees: Functioning in the 1990s*. New York, NY: Ernst & Young.

Francis, J. R., E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (Fall): 17–34.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement): 71–105.

Geiger, M. A., and K. Raghunandan. 2002. Auditor tenure and audit reporting failures. *Auditing: A Journal of Practice & Theory* 21 (March): 67–78.

Hermanson, H. M. 2000. An analysis of the demand for reporting on internal control. *Accounting Horizons* 14 (September): 325–341.

Johnson, V. E., I. K. Khurana, and J. K. Reynolds. 2002. Audit-firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19 (Winter): 637–660.

Kinney, W. R., Jr., and L. S. McDaniel. 1989. Characteristics of firms correcting previously reported quarterly earnings. *Journal of Accounting and Economics* 11 (February): 71–93.

———. 2000. Research opportunities in internal control quality and quality assurance. *Auditing: A Journal of Practice & Theory* 19 (Supplement): 83–90.

Klein, A. 2002a. Economic determinants of audit committee independence. *The Accounting Review* 77 (April): 435–452.

———. 2002b. Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics* 33 (August): 375–400.

KPMG. 1999. *Shaping the Audit Committee Agenda.* Montvale, NJ: KPMG.

Krishnan, G. V. 2003. Audit quality and pricing of discretionary accruals. *Auditing: A Journal of Practice & Theory* 22 (March): 109–126.

Krishnan, J., and J. Krishnan. 1997. Litigation risk and auditor resignations. *The Accounting Review* 72 (October): 539–560.

Krishnan, J. 2002. The timing and information content of auditors' exhibit letters relating to auditor changes. *Auditing: A Journal of Practice & Theory* 21 (March): 29–46.

Lublin, J. S. 2002. Enron audit panel is scrutinized for its cozy ties with the firm. *Wall Street Journal* (February 1).

McDaniel, L., R. D. Martin, and L. A. Maines. 2002. Evaluating financial reporting quality: The effects of financial expertise vs. financial literacy. *The Accounting Review* 77 (Supplement): 139–167.

Myers, J. N., L. A. Myers, and T. C. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (July): 779–799.

———, ———, Z-V. Palmrose, and S. Scholz. 2003. Mandatory auditor rotation: Evidence from restatements. Working paper, University of Illinois at Urbana–Champaign, University of Southern California, and University of Kansas.

National Commission on Fraudulent Financial Reporting (NCFFR, the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: Government Printing Office.

Noreen, E. W. 1989. *Computer Intensive Methods for Testing Hypotheses: An Introduction.* New York, NY: John Wiley & Sons.

Palmrose, Z-V., and S. Scholz. 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (Spring): 139–180.

Petroni, K., and M. Beasley. 1996. Errors in accounting estimates and their relation to audit firm type. *Journal of Accounting Research* 34 (Spring): 151–171.

PricewaterhouseCoopers. 2000. *Audit Committee Effectiveness—What Works Best.* Institute of Internal Auditors Research Foundation (IIARF). Altamonte Springs, FL: IIARF.

Raghunandan, K., and D. V. Rama. 1999. Auditor resignations and the market for audit services. *Auditing: A Journal of Practice & Theory* 18 (Spring): 124–134.

———, W. J. Read, and D. V. Rama. 2001. Audit committee composition, "gray directors," and interaction with internal auditing. *Accounting Horizons* 15 (June): 105–118.

Rittenberg, L. E., and R. D. Nair. 1993. Audit committees: Is there an expectations gap? An analysis of SAS No. 61. *The Expectation Gap Standards: Progress, Implementation Issues, and Research Opportunities*. New York, NY: AICPA.

Securities and Exchange Commission (SEC). 1979. *Statement of Management on Internal Accounting Control*. Release No. 15772 (April), SEC Docket 421-435. Washington, D.C.: SEC.

———. 1988. *Disclosure Amendments to Regulation S-K, Form 8-K and Schedule 14A Regarding Changes in Accountants and Potential Opinion Shopping Situations*. Financial Reporting Release No. 31 (April), SEC Docket 1140-1147. Washington, D.C.: SEC.

———. 1999. *NYSE Rulemaking: Order Approving Proposed Rule Change Amending the Audit Committee Requirements and Notice of Filing and Order Granting Accelerated Approval of Amendments No. 1 and No. 2 Thereto*. Release No. 34-42233. Available at: http://www.sec.gov/rules/sro/ny9939o.htm.

———. 2002. *Final Rule: Certification of Disclosure in Companies' Quarterly and Annual Reports*. Release Nos. 33-8124; 34-46427. Available at: http://www.sec.gov/rules/final/33-8124.htm.

———. 2003a. *Final Rule: Management's Reports on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports*. Release Nos. 33-8238; 34-47986. Available at: http://www.sec.gov/rules/final/33-8238.htm.

———. 2003b. *Final Rule: Strengthening the Commission's Requirements Regarding Auditor Independence*. Release Nos. 33-8183; 34-47265; 35-27642. Available at: http://www.sec.gov/rules/final/33-8183.htm.

———. 2003c. *Final Rule: Disclosure Required by Sections 406 and 407 of the Sarbanes-Oxley Act of 2002*. Release Nos. 33-8177; 34-47235. Available at: http://www.sec.gov/rules/final/33-8177.htm.

Stice, J. D. 1991. Using financial and market information to identify pre-engagement factors associated with lawsuits against auditors. *The Accounting Review* 66 (October) 516–533.

U.S. House of Representatives, Committee on Financial Services. 2002. *Sarbanes-Oxley Act of 2002*. Public Law No. 107-204. Washington, D.C.: Government Printing Office.

Verschoor, C. C. 2002. Reflections on the audit committee's role. *Internal Auditor* 59 (April): 26–35.

Whitehead, J. C. 2002. Prepared Statement of Mr. John C. Whitehead Former Co-Chairman of Goldman Sachs & Co. at Hearing on Accounting and Investor Protection Issues Raised by Enron and Other Public Companies. Available at: http://banking.senate.gov/02_03hrg/031902/whithead.htm.

Wright, A., and S. Wright. 1996. The relationship between assessments of internal control strength and error occurrence, impact and cause. *Accounting and Business Research* 27 (Winter): 58–71.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–82.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 677–702

Profit Sharing in an Auditing Oligopoly

Xiaohong Liu
Hong Kong University of Science & Technology

Dan A. Simunic
University of British Columbia

ABSTRACT: This paper examines how partners in an audit firm can use profit-sharing rules to induce optimal partner behavior from the firm's point of view, taking into account the strategic competition of firms in an auditing oligopoly. We use a linear contracting framework to investigate the effects of profit-sharing rules on individual partners' various decisions, including their pricing strategies and effort choices. We assume that efficient audits of different types of clients require different effort profiles with respect to degree of partner cooperation. For example, the audit of a complex company requires different amounts of partner collaboration than does the audit of a simple company. Moreover, since it is too costly for an enforcement party, such as the head office of an audit firm or a court, to verify each client's type in order to resolve compensation disputes among the firm's partners, it is reasonable to assume that client type cannot be contracted upon for partner compensation purposes. Given this assumption, we derive conditions under which there exists an equilibrium in which audit firms strategically choose different profit-sharing rules to specialize in different types of clients, thereby earning positive economic profits. Our analysis provides insights into the strategic competition among the big audit firms, and helps to explain the observed differences in the compensation plans of these firms and in the nature of their client portfolios.

Keywords: *audit partnership; profit-sharing rules; oligopoly; specialization.*

I. INTRODUCTION

In this study, we investigate incentives and profit-sharing rules in audit firms and their implications for the audit services market. Existing literature normally treats an audit firm as a single person "auditor," and pays limited attention to the decentralized nature

---

We thank two anonymous reviewers, and workshop participants at the University of British Columbia, University of Waterloo, University of Toronto, University of Wisconsin–Madison, and Hong Kong University of Science & Technology for their comments. Professor Liu is very grateful to Jerry Feltham for his valuable advice and suggestions. We also thank the partners from Arthur Andersen & Co. and KPMG for providing us with information about compensation plans in their firms. Finally, we thank the Wei Lun Fellowship and the Social Sciences & Humanities Research Council of Canada for financial support. This paper is based on Professor Liu's Ph.D. thesis at the University of British Columbia.

Editor's note: This paper was accepted by Madhav Rajan, Editor.

Submitted June 2003
Accepted July 2004

of actual firms in which individual auditors operate as partners.[1] Like other rational players in the economy, audit partners make decisions to maximize their own economic interests, which are influenced by the reward and control systems of the partnership. Hence, understanding the profit-sharing rules of audit firms is crucial to understanding auditors' behavior.[2] As Zeff (2003) has commented, the manner in which audit firms reward their partners is a fundamental issue facing the accounting profession.

Within an audit firm, partners make decentralized decisions in performing various tasks—including bidding for clients, expending effort in performing audits, and assisting one another in the performance of audits. A moral hazard problem (Holmstrom 1982) arises when an individual partner takes actions that are inconsistent with the interests of the partnership as a whole. For example, if the legal liability costs associated with an audit are shared equally among all partners in the firm, then the engagement partner's effort level can be expected to fall short of the optimal level from the firm's perspective, because the engagement partner does not fully bear the consequences of his effort choice. A principal objective of this study is to examine how partners in an audit firm can use profit-sharing rules that base an individual partner's compensation on his performance—as reflected by audit fee revenues and liability costs from lawsuits—to induce optimal partner behavior from the firm's point of view.

Audit firms are normally divided into two groups based on brand name or firm size: the Big 4 versus the non-Big 4, with the Big 4 firms dominating the audit services market.[3] For example, 73 percent of the 7,006 publicly traded companies included in *Who Audits America* (2003) were audited by the Big 4, and the rest were audited by 81 non-Big 4 firms. With just a few firms dominating the market, it is reasonable to expect these firms to act as strategic competitors. In this study, we examine how audit firms can use their profit-sharing rules as strategic instruments to compete in an oligopolistic audit market.

A number of studies have documented differences in the profit-sharing rules used by the big audit firms. For example, based on survey responses in the late 1980s and early 1990s, Trompeter (1988, 1994) classifies the compensation plans of the then Big 8 (6) firms into "small pool" and "large pool" plans, according to the importance of "local office profitability" in determining an individual partner's compensation.[4] He reports:

> For Local Office Profitability, there is an interesting dichotomy in the [partners'] responses. Office managing partners from three of the firms deemed this component to be unimportant in determining partner compensation. These firms base partner compensation on firm-wide profitability as opposed to office-wide profitability. Partners from the other firms rated this factor very important. (Trompeter 1988, 85)

Using the "large-pool" versus "small-pool" classification, Trompeter (1994) and Carcello et al. (2000) investigate the effects of different profit-sharing arrangements on partners' reporting decisions, hypothesizing that partners from "small pool" firms are more likely to compromise their objectivity and independence than are those from "large pool" firms. They find some evidence consistent with this prediction.

---

[1] An exception is Narayanan (1995), but he treats profit-sharing rules as exogenous in the model.

[2] The terms "partner" and "auditor" are used interchangeably in this paper.

[3] The Big 4 are Deloitte Touche Tohmatsu, Ernst & Young, KPMG, and PricewaterhouseCoopers. The number of big audit firms shrank from eight to five through mergers in the 1980s and 1990s, and to four after the elimination of Arthur Andersen from the group as a result of the Enron debacle in 2001.

[4] Burrows and Black (1998) provide similar observations. In particular, they report that the partners in the Melbourne office of Arthur Andersen (KPMG, respectively) share in the world wide profit pool (share only in the Melbourne office profit pool).

While interesting, these studies do not consider certain fundamental questions: Why do the big audit firms use different compensation schemes? How do these differences affect partners' decisions in performing multiple tasks and the outcome of competition? By addressing these questions, we seek to shed light on the strategic competition among the big audit firms and to provide an explanation for the observed differences in the compensation plans of these firms.

In our analysis, we use a stylized model consisting of two audit firms, with two auditors (namely partners) in each firm, operating separately in offices located in two geographically different places. Audit production is viewed as a two-stage process. In the first stage, auditors from different firms in each location compete in bidding for clients. In the second stage, auditors in each firm cooperate in performing the audits for the clients obtained in the first stage.

We assume that the efficient performance of an audit requires some degree of teamwork by partners (from different offices) within a firm. This assumption is consistent with the notion that the economic benefits of teamwork are the main reason for auditors to join one another in forming a partnership. In fact, one can reasonably argue that the big firms exist as multinational partnerships in order to facilitate collaboration among their offices (and partners) so as to more efficiently service multinational client companies.[5] The potential benefits of collaboration may also occur when a client company operates in multiple industry sectors. Ferguson et al. (2003) report evidence that industry specialization is a characteristic of specific big firm offices and partners—not the firm as a whole. Thus, the collaboration of industry-specialist firm offices (partners) may be needed to efficiently audit a client that operates in multiple industry sectors.

While partner cooperation can be useful in many contexts, the amount of partner cooperation needed for an audit will likely vary across clients. Client companies differ greatly in size, geographic dispersion and complexity of their operations, extent of their operations in multiple industries, and other characteristics that can influence the degree of partner collaboration needed to produce an audit efficiently. For example, the appropriate level of partner collaboration may depend on the degree of industry-specific knowledge required for the audit, and the audit of a large international company probably requires more partner cooperation than the audit of a small local company. In our analysis, we capture such differences among audit clients by assuming that clients can be classified into two types: complex clients and simple clients.

Turning to results, we find that, given that an efficient audit requires some degree of teamwork, it is optimal for partners within a firm not only to share the legal liability costs of audits, but also to share engagement revenues (i.e., audit fee revenues). Liability sharing between partners is essential for the motivation of teamwork because an auditor is willing to provide consultation/help to his partner if and only if he bears some responsibility for the audit failure of his partner's clients. To understand the need for revenue sharing, we observe that a partner would price below the firm's cost in bidding for clients if he received the whole amount of his engagement revenues. This occurs because the partner's personal cost of an audit is less than the firm's cost (since he bears only a part of the legal liability cost of the audit and ignores his colleague's cooperation/helping effort). Consequently, partners should share each other's engagement revenues to counteract their incentives to bid below the firm's cost.

---

[5] See, e.g., Cypert (1991), who gives an account of how the benefits of cross-justification cooperation of audit partners drove the mergers that formed KPMG.

Moreover, we find that firms' equilibrium compensation strategies differ significantly, depending on whether client type is contractable or is not contractable. When client type is contractable, firms have the same dominant strategy, consisting of client-type-specific sharing rules, which induce optimal partner behavior (in terms of effort profiles and pricing policies) for each type of client. This is a perfect competition equilibrium in which auditors from different firms provide audits of the same value, bid the same price, and earn zero economic profit.[6]

However, we believe that, in the real world, client type is unlikely to be contractable because a client's operating characteristics that determine its "type" (i.e., complex versus simple and the resulting effort mix for an efficient audit) are generally difficult to verify. This seeming contracting disadvantage may, however, create an economic advantage for audit firms. We find that, when client type is not contractable and under suitable conditions, there exists a specializing equilibrium in which firms strategically use different compensation schemes to induce efficient audits for different types of clients, thereby gaining limited monopoly power over the clients in which they specialize. In equilibrium, each firm provides higher value audits for clients in its niche market than does the rival firm, and earns positive economic profits. These equilibrium results can explain the existence of both "large profit pool" and "small profit pool" big firms observed by Trompeter (1988, 1994) and can also help explain client-auditor alignments. For example, our theory is consistent with the observation that the client portfolio of Arthur Andersen, a former "large pool firm," was concentrated in large, complex companies, while KPMG, a "small pool firm," was concentrated in smaller, simpler clients (U.S. Senate 1977; Liu 2002). We provide more discussion of these observations and their links to our analytical results in Section V.

The remainder of the paper is organized as follows: Section II presents the model; Sections III and IV derive the equilibrium under the assumptions that client type is contractable and not contractable, respectively; Section V concludes the paper. A table of notation and all proofs are given in the Appendix.

II. THE MODEL

The Shareholders' Problem

Consider a one-period model with two geographically different locations and two audit firms. All parties in the model are assumed to be risk-neutral and the interest rate is normalized to zero. Each audit firm has two partners operating separately in the two locations. This assumption captures a real-world situation in which an audit firm establishes offices in different states (and countries) with partners operating separately in different offices. A partner is assumed not to solicit clients outside his jurisdiction (namely location). We label the two locations 1 and 2, and the two firms A and B, so that Auditor A1, for example, refers to the partner of Firm A who operates in Location 1 (see Figure 1). The number of companies available for audit in each location is denoted by a positive integer n_i, $i \in \{1, 2\}$.

---

[6] We distinguish between the terms "accounting profit" and "economic profit." Accounting profit is the difference between the audit fee revenues and the liability costs, whereas economic profit equals accounting profit minus the auditors' effort costs. Unlike accounting profit, economic profit is not verifiable because only individual auditors have private information about their effort choices. In this paper, the term "profit" refers to accounting profit, unless otherwise noted.

FIGURE 1
Auditors from Firms A and B in Two Geographically Different Locations

Partner A1 Partner B1	Partner A2 Partner B2
Location 1	Location 2

At the beginning of the period, shareholders decide whether to continue investing in or liquidate their companies. *A priori*, a company is either a good company with probability $\phi \in (0, 1)$ or a bad company with probability $(1 - \phi)$, depending on its financial condition, which is not known to the company's shareholders. A good company is worth $V > 0$ if liquidated now, or returns $V' > V$ at the end of the period if the investment is continued, whereas a bad company is worth $I > 0$ if liquidated now, or returns zero if the investment is continued. We assume that the expected return from continuing investment in a company is greater than that from immediate liquidation of the company, i.e.:

$$\phi V' > \phi V + (1 - \phi)I. \tag{1}$$

Thus, shareholders will continue investing in a company based on their prior beliefs, and their (expected) payoff from the company is:

$$\phi V'. \tag{2}$$

The shareholders may engage an audit firm to perform an audit of the company's financial condition. To focus on the audit firm's problem, as is common in the literature (e.g., Dye 1995; Schwartz 1997), we assume that the audit firm is the sole preparer of an attestation report on the company's financial condition.[7] Once the audit firm is engaged, the shareholders pay an audit fee, which is determined before the audit is performed (i.e., a fixed audit fee contract).[8]

As discussed in Section I, we assume that the performance of an audit requires some degree of teamwork in the sense that the engagement partner's (that is, the audit firm's) attestation report is the outcome of partners' joint effort inputs in the audit process. We assume that a partner's effort can be measured in terms of personal costs, and that the partner's effort is additively separable across tasks and clients. Formally, for each client, let $e \in [0, \bar{e}]$ be the cost of the audit effort by the engagement auditor (partner), and

---

[7] In the real world, the role of an auditor is to render an opinion on whether the financial reports prepared by the company's manager fairly present the financial condition of the company. To simplify the analysis and to focus on the audit firm's problem, in the model we do not explicitly consider the agency problem and contracting issues between shareholders and managers, and assume that auditors prepare the financial reports.

[8] This assumption is consistent with most real-world situations. For example, a 2002 survey (1999 survey, respectively) by the Committee on Corporate Reporting found that 21 (30) out of the 25 (31) big audit firms' clients surveyed engaged their auditors with annual fixed fee contracts (see http://www.fei.org/download/2002_Audit_Fee.xls).

$h \in [0, \bar{h}]$ the cost of the cooperation/helping effort by his partner. The audit technology is characterized by the probabilities of the engagement auditor's correct attestation report conditional on the company's real financial condition and the partners' joint effort inputs as follows:

$$\Pr[g|\text{Good}, e, h] = 1 \text{ and } \Pr[b|\text{Bad}, e, h] = \mu(e, h) \tag{3}$$

where g (b) denotes the engagement auditor's report that the company is good (bad). That is, we assume that there is no Type I error in the sense that the engagement auditor will not issue a "b" report on a good company; on the other hand, the auditor may issue a "g" report on a bad company.[9] With such a specification of audit technology, we represent the audit quality by μ, namely the conditional probability of detecting a bad company.

As discussed in Section I, the audits of different types of client companies require different amounts of partner collaboration, and client types are determined by companies' operating characteristics. To model this situation in a tractable manner, we assume two types of clients, complex and simple, and denote them by two real numbers C and S, respectively, where $C < S$. We let the audit quality function $\mu(\cdot)$ be parameterized by a real number $x \in (\underline{x}, \bar{x})$, and let the set of client types be a subset of the parameter set. That is:

$$\mu\colon [0, \bar{e}] \times [0, \bar{h}] \times (\underline{x}, \bar{x}) \rightarrow [0, 1] \text{ and } \{C, S\} \subset (\underline{x}, \bar{x}). \tag{4}$$

We represent the numbers of different types of clients in each location by positive integers n_i^t, where $i \in \{1, 2\}$ and $t \in \{C, S\}$. Note that $n_i^C + n_i^S = n_i$.

We assume that the audit quality function, $\mu(e, h; x)$, is continuously differentiable and has the following properties on its domain, with the subscripts e, h, and x denoting derivatives:

Assumption 1: $\mu = 0$ if $e = 0$; $\mu = 1$ if, and only if, $(e, h) = (\bar{e}, \bar{h})$.

Assumption 2: $\mu_e > 0$; $\mu_h \geq 0$, with strict inequality if $e > 0$.

Assumption 3: $\mu_{ex} \geq 0$; $\mu_{hx} \leq 0$, with strict inequality if $e > 0$.

Assumption 4: $\mu_{ee} < 0$ and $\mu_{ee}\mu_{hh} - \mu_{eh}\mu_{eh} > 0$ for any (e, h) in the set $(0, \bar{e}) \times (0, \bar{h})$.

Assumption 5: $\mu_{eh} > 0$.

The first part of Assumption 1 means that the engagement auditor's effort is essential, in that the probability of detecting a bad company is zero if the engagement auditor puts forth zero effort (regardless of his partner's cooperative effort). The second part of Assumption 1 means that a perfect audit is conducted (i.e., the probability of detection is 1) if and only if both partners' effort levels attain the upper bounds. Assumption 2 means that audit quality increases with partners' effort.[10] Assumption 3 implies that the marginal benefit of engagement audit effort e (cooperation effort h) increases (decreases) with client type. Assumption 4 implies that μ is a strictly concave function of effort pair (e, h), which means that the marginal benefit of a person's effort is decreasing with his own effort. Assumption 5 states that the marginal benefit of a person's effort is increasing with his partner's effort; i.e., the partners' effort inputs are complements in the audit process. By way of an example, note that these assumptions are satisfied by the following class of functions: $\mu(e, h; x, \rho) = xe^{1/\rho} + (1 - x)(eh)^{1/\rho}$, where $e \in [0, 1]$, $h \in [0, 1]$, $x \in (0, 1)$, and $\rho \geq 2$.

---

[9] These assumptions of audit technology are reasonable since they translate into the real-world situations that (1) an auditor will not find any misstatement in the manager's financial report when there is no misstatement and (2) an auditor may fail to detect material misstatements in the manager's financial report when the report is misstated. These assumptions of audit technology are also common in the literature; see, e.g., Dye (1995) and Pae and Yoo (2001).

With the audit technology specified in Equation (3), it is clear that audit failure (conditional on a company being bad) occurs with probability $(1 - \mu)$. For simplicity, we consider a strict liability regime in which the audit firm is liable whenever shareholders suffer losses as a result of audit failure.[11] That is, when audit failure occurs and a bad company continues to operate, the shareholders suffer financial losses—which equal their investments in the bad company, I, in the context of this model—and get legal payments, denoted by $l > 0$, from the audit firm through litigation.

We now examine the shareholders' decision problem. If the audit firm issues a "b" report, then the shareholders will liquidate the company since there is no Type I error (i.e., no incorrect "b" report). On the other hand, if the audit firm issues a "g" report, then the shareholders will continue their investments in the company. The reason is as follows. Given a "g" report, the shareholders' payoff from continuation of their investments is:

$$\Pr[\text{Good}|g]V' + \Pr[\text{Bad}|g]l, \tag{5}$$

where $\Pr[\text{Good}|g] = \dfrac{\phi}{\phi + (1 - \phi)(1 - \mu)}$, and $\Pr[\text{Bad}|g] = 1 - \Pr[\text{Good}|g]$. In contrast, given a "$g$" report, the shareholders' payoff from liquidation is:

$$\Pr[\text{Good}|g]V + \Pr[\text{Bad}|g]I. \tag{6}$$

By comparing Expressions (5) and (6), we know that shareholders given a "g" report get a higher payoff from continuation than from liquidation, because:

$$\begin{aligned}\Pr[\text{Good}|g]V' + \Pr[\text{Bad}|g]l &= \frac{\phi V' + (1 - \phi)(1 - \mu)l}{\phi + (1 - \phi)(1 - \mu)}\\ &> \frac{\phi V + (1 - \phi)I + (1 - \phi)(1 - \mu)l}{\phi + (1 - \phi)(1 - \mu)}\\ &> \frac{\phi V + (1 - \phi)(1 - \mu)I}{\phi + (1 - \phi)(1 - \mu)}\\ &= \Pr[\text{Good}|g]V + \Pr[\text{Bad}|g]I,\end{aligned}$$

---

[10] Assumptions 1 and 2 are similar to those in Itoh (1991), who models mutual help between agents in a principal-agent setting.

[11] We considered a negligence liability regime in an earlier version of the paper, but that merely complicates the notation without adding any new insights. See Schwartz (1997) for a discussion of different legal regimes and their implications.

where the first inequality comes from substitution from Inequality (1). Based on the shareholders' optimal continuation/liquidation decisions conditional on different audit reports, the shareholders' *ex ante* payoff from purchasing an audit is:

$$\begin{aligned}&\Pr[g](\Pr[\text{Good}|g]V' + \Pr[\text{Bad}|g]l) + \Pr[b]\Pr[\text{Bad}|b]I - f\\&\quad= \phi V' + (1 - \phi)(1 - \mu)l + (1 - \phi)\mu I - f,\end{aligned} \tag{7}$$

where f denotes the audit fee.

Taking the difference between the shareholders' *ex ante* payoffs with and without an audit, as stated in Equations (7) and (2), the shareholder surplus (i.e., the shareholders' net benefit) from purchasing an audit is:

$$(1 - \phi)\mu I + (1 - \phi)(1 - \mu)l - f, \tag{8}$$

which represents the expected savings of the wasteful investments in a bad company plus the expected legal payments from the audit firm through litigation, less the audit fee. The shareholders' problem in choosing an audit firm is to maximize the shareholder surplus.

The Audit Firm's Problem

Since audit technology depends on client type, as shown in Condition (4), an audit firm may wish to vary its compensation structure according to client type in order to induce different partner behavior suitable for different types of clients. However, realistically, it would be too costly, if not impossible, for an enforcement party (such as the head office of the firm or a court) to verify each client's type in order to settle compensation disputes among the firm's partners. This is because the client's operating characteristics, which determine both its type and the necessary amount of partner cooperation for an efficient audit, are generally difficult to verify. Even if some of the information associated with client type is verifiable—such as the number of subsidiaries a client has in different locations, or the number of industry sectors in which it operates—this information may not perfectly classify a client. To see this, consider a case in which a company has many subsidiaries located in different countries. If the affiliated companies engage in many intercompany transactions, then they are likely to be complex clients because a significant degree of cross-jurisdiction partner consultation may be needed to address difficult auditing issues arising from transfer-pricing, tax-planning, and related-party transactions. However, if the affiliated companies' operations are relatively independent, then they are likely to be simple clients, and their audits may not require much partner collaboration.

The emphasis of our analysis, then, is not on the case in which client type is contractable, but on the more realistic case in which client type is not contractable. Nevertheless, since the analysis of the latter can be built on the former, we analyze both cases in order to better illustrate their differences.

III. CLIENT TYPE CONTRACTABLE

Profit-Sharing Rules

An audit firm's profit-sharing rules consist of revenue-sharing and liability-sharing rules, which are not necessarily the same. For tractability of the analysis, the sharing rules are assumed to be linear:

$$y_i = \sum_{t\in\{C,S\}} \alpha_i^t R_i^t + \sum_{t\in\{C,S\}} (1-\alpha_j^t) R_j^t - \sum_{t\in\{C,S\}} \beta_i^t L_i^t - \sum_{t\in\{C,S\}} (1-\beta_j^t) L_j^t + \eta_{ij}, \qquad (9)$$

where:

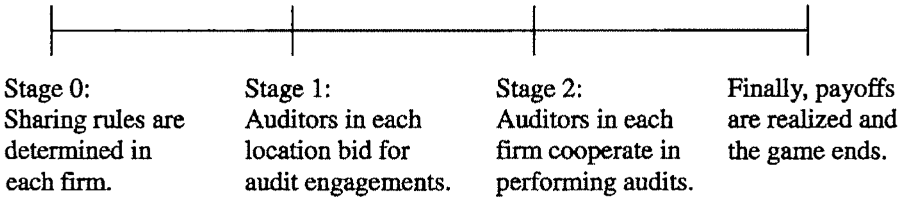
y_i = Partner i's compensation;
i, j = partners in the firm, $i \neq j$, $i, j \in \{1, 2\}$;
t = the client's type, $t \in \{C, S\}$;
$R_i^t \geq 0$ = audit fee revenues related to the t type clients audited by Partner i;
$L_i^t \geq 0$ = legal liability costs related to the t type clients audited by Partner i;
α_i^t = finite incentive weight on revenues related to t type clients audited by Partner i;
β_i^t = finite incentive weight on legal liability costs related to t type clients audited by Partner i; and
η_{ij} = lump sum transfer between partners, $\eta_{12} = -\eta_{21}$.

Accordingly, the revenue-sharing (liability-sharing) rules, characterized by α (β), specify how the partners share each other's engagement revenues (legal liability costs) within the firm.

Timeline for the Auditors' Game

As depicted in Figure 2, the timeline for the auditors' game is as follows: At Stage 0, *in each firm*, partners choose their profit-sharing rules. We assume that a partner agrees to participate in the game if his expected payoff given the profit-sharing rules is non-negative. At Stage 1, *in each location*, auditors from different firms submit bids, under the names of their firms, to shareholders in their jurisdiction. For each client, the auditor who provides shareholders with the higher surplus wins the bid. At Stage 2, *in each firm*, partners choose effort levels in performing audits and in helping each other in the performance of audits. Finally, companies' returns are realized and thus the companies' natures (i.e., whether they are good or bad companies) are revealed; legal payments are made by audit firms to shareholders if shareholders suffer losses as a result of audit failure; partners are compensated according to the sharing rules specified at Stage 0. As noted earlier (see footnote 2), the terms "partner" and "auditor" are used interchangeably in this paper, although sometimes we use one term over the other in order to emphasize one of the person's dual roles: an auditor, who performs audits for shareholders, is also a partner who provides cooperation/helping effort to his colleague and makes various decisions within the audit firm.

FIGURE 2
Timeline for the Auditors' Game



Stage 0:
Sharing rules are determined in each firm.

Stage 1:
Auditors in each location bid for audit engagements.

Stage 2:
Auditors in each firm cooperate in performing audits.

Finally, payoffs are realized and the game ends.

This is a three-stage game with observed actions.[12] In particular, we assume that the firms' sharing rules become observable to all the players before the start of Stage 1. We believe this assumption is a reasonable description of the real world. For example, Burrows and Black (1998, 522) interviewed partners from the Melbourne offices of the then Big 6 firms and report that "despite the dearth of published information on the topic, each partner knew in general how rival firms shared profits."[13] This assumption is also commonplace in the literature on strategic delegation in oligopoly firms (e.g., Fershtman et al. 1991; Reitman 1993; Aggarwal and Samwick 1999). As Fershtman and Judd (1987) argue, firms' compensation contracts are infrequently altered and repeated play of the game should enable interested parties to learn about these contracts. We therefore solve the game for a subgame perfect equilibrium through backward induction.

Stage 2's Problem: Effort Choice in the Performance of Audits

The effort choice subgame in Stage 2 consists of two independent, simultaneous games, each played by partners within an audit firm. Given symmetry, it suffices to analyze the problem of just one firm. Within the firm, for an arbitrarily chosen client, each partner's objective is to choose an effort level to minimize his personal costs (comprised of his effort cost and *his share* of the expected legal liability costs associated with the audit), treating sharing rules and his belief about his partner's effort choice as given. Formally, with $\mu(e, h; t)$ written more concisely as $\mu^t(e, h)$, the engagement auditor's decision problem is:

$$\min_{e\in[0,\bar{e}]} \beta_i^t (1 - \phi)(1 - \mu^t(e, h))l + e, \tag{10}$$

and his partner's decision problem for this client is:

$$\min_{h\in[0,\bar{h}]} (1 - \beta_i^t)(1 - \phi)(1 - \mu^t(e, h))l + h, \tag{11}$$

where $i \in \{1, 2\}$ and $t \in \{C, S\}$. Observe in Objective Functions (10) and (11) that the revenue-sharing rule, α, does not enter into partners' effort decisions. This is because, given fixed-audit-fee contracts, the revenues a partner receives are fixed when he makes effort decisions.

We restrict the liability-sharing rule, β, to non-negative values, since the firm can always replace a negative value for β with zero to induce the same effort profile (that is, zero effort in the audit). Furthermore, we impose the following technical assumption to ensure an interior solution:

Assumption 6: $\lim_{e\to 0} \mu_e \to \infty$, $\lim_{h\to 0} \mu_h \to \infty$ provided $e > 0$; $(1 - \phi)\mu_e(\bar{e})l < 1$, $(1 - \phi)\mu_h(\bar{h})l < 1$.

Proposition 1: In an audit firm, given the liability-sharing rules characterized by $\{\beta_1^C, \beta_1^S, \beta_2^C, \beta_2^S\}$, a partner chooses the same unique level of effort—audit effort e or cooperation/helping effort h—for every client of the same type. Specifically, if $\beta_i^t = 0$, then $e_i^{t*} = h_j^{t*} = 0$, where $i \neq j$,

---

[12] The assumption that partners observe audit fees before choosing effort is not essential because, given fixed audit fee contracts, effort choices do not depend on the audit fees set in Stage 1.

[13] Our interviews with the partners from Arthur Andersen and KPMG also suggest that audit partners are knowledgeable about the rival firms' compensation schemes.

$i, j \in \{1, 2\}$, $t \in \{C, S\}$. If $\beta_i^t \geq 1$, then $e_i^{t*} \in (0, \bar{e}]$ and $h_j^{t*} = 0$. If $\beta_i^t \in (0, 1)$, then $e_i^{t*} \in (0, \bar{e})$ and $h_j^{t*} \in (0, \bar{h})$, and (e_i^{t*}, h_j^{t*}) are determined by the following conditions:

$$\beta_i^t(1 - \phi)\mu_e^t(e, h)l - 1 = 0, \tag{12}$$

$$(1 - \beta_i^t)(1 - \phi)\mu_h^t(e, h)l - 1 = 0. \tag{13}$$

Proposition 1 shows that the equilibrium effort profile for an audit, (e_i^{t*}, h_j^{t*}), is determined by the firm's liability-sharing rule, β, and by the client's type, t. First, observe that an auditor has no incentive to help his partner if he bears no responsibility for his partner's audit failure (i.e., $h_j^{t*} = 0$ if $\beta_i^t \geq 1$), and the auditor is induced to provide a positive amount of cooperation/helping effort if he bears some responsibility for his partner's audit failure (i.e., $h_j^{t*} > 0$ if $0 < \beta_i^t < 1$).

Next, observe that the same liability-sharing rule induces different effort profiles for different types of clients. More specifically, given any liability-sharing rule $\beta \in (0, 1)$, the equilibrium effort level changes with client type, with the direction of the change dependent on two opposing forces. On the one hand, based on the Equilibrium Conditions (12) and (13), Assumption 3 implies that when the client type increases from C to S, the direct effects are that audit effort e (cooperation effort h) increases (decreases) for a given level of the other partner's effort. On the other hand, Assumption 5 implies that partners' effort inputs are strategic complements in the sense that the increase (decrease) of one partner's effort raises (reduces) the other partner's incentive to increase effort. Therefore, the overall effect of the increase of client type on effort depends on which of the two forces dominates.[14]

Stage 1's Problem: Pricing Strategies

The subgame in Stage 1 consists of two independent, simultaneous games played in two locations, and in each location, auditors from different firms compete in bidding for clients. Given symmetry, we focus on auditors' bidding/pricing strategies in just one location, say, Location $i \in \{1, 2\}$. We first characterize the competing auditors' payoffs from an audit engagement. We let κ_{ri}^t represent Auditor ri's personal cost of the audit, and K_{ri}^t represent the firm's cost (i.e., the cost to the firm's partners as a whole), where $r \in \{A, B\}$ and $t \in \{C, S\}$. Based on Proposition 1, we have:

$$\kappa_{ri}^t \equiv \kappa^t(\beta_{ri}^t) = \beta_{ri}^t(1 - \phi)[1 - \mu^t(e^{t*}(\beta_{ri}^t), h^{t*}(\beta_{ri}^t))]l + e^{t*}(\beta_{ri}^t),$$

$$K_{ri}^t \equiv K^t(\beta_{ri}^t) = (1 - \phi)[1 - \mu^t(e^{t*}(\beta_{ri}^t), h^{t*}(\beta_{ri}^t))]l + e^{t*}(\beta_{ri}^t) + h^{t*}(\beta_{ri}^t).$$

Hence, Auditor ri gains a personal payoff of $(\alpha_{ri}^t f_{ri}^t - \kappa_{ri}^t)$ if he wins the audit engagement, whereas the firm's payoff/economic profit from the audit is $(f_{ri}^t - K_{ri}^t)$ (see footnote 6 for the definition of the firm's economic profit). Note that the personal payoff may differ from the firm's payoff because, when the personal payoff is calculated, the self-interested auditor discounts the audit fee revenue and the expected legal liability costs associated with the audit by α and β, respectively, and completely ignores his partner's cost of cooperation effort, h. The incentive problem caused by the discrepancy between the auditor's personal payoff and the firm's payoff will soon become clear (see discussion following Proposition 2).

---

[14] In particular, it can be shown by Cramer's rule and Assumptions 3–5 that, given any $\beta \in (0,1)$, the equilibrium audit effort e^* (cooperation effort h^*) strictly increases (decreases) as client type changes from C to S, when the direct effect dominates, i.e., when $\mu_{ex} > (\mu_{hx}/\mu_{hh})\mu_{eh}$ (when $-\mu_{hx} > (-\mu_{ex}/\mu_{ee})\mu_{eh}$, respectively) for all $x \in [C, S]$.

We now analyze the auditors' bidding strategies. Recall from the analysis in Section II that shareholders will choose the auditor who provides the higher shareholder surplus, as stated in Expression (8), which equals the gross benefit of the audit, denoted by $v^t_{ri} \equiv v^t(\beta^t_{ri}) = (1 - \phi)\mu^t_{ri}I + (1 - \phi)(1 - \mu^t_{ri})l$, less the audit fee, f^t_{ri}. Since the gross benefit of an audit is fixed given the firm's liability-sharing rule, an auditor's choice variable in this stage is the audit fee. In bidding for clients, each auditor's objective is to choose an audit fee so as to provide just a slightly higher shareholder surplus than does his competitor, while not allowing the audit fee to fall below his break-even price, which is denoted by $\underline{f}$, and $\underline{f}^t_{ri} \equiv \kappa^t_{ri}/\alpha^t_{ri}$ by definition. We let σ^t_{ri} be the maximum amount of shareholder surplus Auditor *ri* can provide to shareholders, which is the difference between the gross benefit of the audit and the auditor's break-even price: $\sigma^t_{ri} \equiv v^t_{ri} - \underline{f}^t_{ri}$. Note that we restrict revenue-sharing rules to positive values (i.e., $\alpha > 0$), because a firm chooses a non-positive α only if it wants to discourage its partners from bidding for clients, but the firm can always use a sufficiently small positive α such that the maximum shareholder surplus is negative (i.e., $\sigma < 0$) to achieve that objective.

As in Grossman (1981) and Hurter and Lederer (1986), we impose the feasibility restriction that an auditor will not bid below his break-even price. In addition, to ensure the existence of equilibrium, we make the following tie-breaking assumption about shareholders' preferences:

Assumption 7: If auditors provide the same positive shareholder surplus, then shareholders choose the auditor for whom the maximum shareholder surplus is higher, and they choose auditors randomly with equal chance if the maximum shareholder surplus is the same for all auditors.

Proposition 2: In Location $i \in \{1, 2\}$, given the firms' sharing rules characterized by $\{\alpha^t_{qi}, \beta^t_{qi}, \alpha^t_{ri}, \beta^t_{ri}\}$, $t \in \{C, S\}$, $q \neq r$, $q, r \in \{A, B\}$, for each client, if the maximum shareholder surplus, σ, is negative for both auditors (i.e., if $\sigma^t_{qi} < 0$ and $\sigma^t_{ri} < 0$), then both auditors have zero demand. If $\sigma^t_{qi} < 0$ and $\sigma^t_{ri} \geq 0$, then the auditor from Firm q has zero demand, and the auditor from Firm r wins the bid with the monopoly price equal to the gross benefit of the audit, v^t_{ri}, and gains a non-negative personal payoff (a positive payoff if $\sigma^t_{ri} > 0$). If $\sigma^t_{qi} = \sigma^t_{ri} \geq 0$, then both auditors bid their break-even prices and win the bid with equal chance. If $0 \leq \sigma^t_{qi} < \sigma^t_{ri}$, then the auditor from Firm q bids his break-even price, $\underline{f}^t_{qi}$, and the auditor from Firm r wins the bid with the price equal to $(v^t_{ri} - \sigma^t_{qi})$ and gains a positive personal payoff.

The bidding/pricing equilibrium characterized by Proposition 2 is intuitive: given that shareholders choose the auditor who provides the higher (positive) shareholder surplus, the auditor for whom the maximum shareholder surplus is higher has a competitive advantage, in the sense that the auditor can always offer a higher shareholder surplus (relative to his competitor) with an audit fee greater than his break-even price. However, as noted earlier, the auditor's personal payoff may differ from the firm's payoff for an audit engagement. Consequently, the auditor's equilibrium pricing strategy may not be optimal from the partnership's point of view. In particular, an auditor has an incentive to bid below the firm's cost when his break-even price, $\underline{f}$, falls below the firm's cost, K. To illustrate, consider a case where $0 \leq \sigma^t_{B1} < \sigma^t_{A1}$. In this case, Auditor A1 wins the bid and gains a positive payoff with the equilibrium price $f^{t*}_{A1} = v^t_{A1} - \sigma^t_{B1} > \underline{f}^t_{A1}$ (by Proposition 2), but Firm A as a whole suffers a loss when the auditor's break-even price is less than the firm's cost, i.e., when $\underline{f}^t_{A1} < f^{t*}_{A1} < \mathrm{K}^t_{A1}$. Conversely, it is easy to see that an auditor has an incentive to forgo a bid profitable to the firm when his break-even price exceeds the firm's cost, i.e., when $\underline{f} > \mathrm{K}$.

Stage 0's Problem: Profit-Sharing Rules

At the contracting stage, observe that within each audit firm both partners prefer the sharing rules that maximize the sum of their payoffs, which is just the firm's payoff, namely the firm's economic profit. Thus, the game in Stage 0 can be viewed as a simultaneous game between the two firms, with each firm choosing its sharing rules to maximize its payoff against the compensation strategy of the rival firm. To focus on the incentive effects of the sharing rules, we will ignore the determination of η, namely the predetermined lump sum transfer between the partners of a firm, because the transfer does not affect the partners' effort or pricing decisions (i.e., it has no incentive effect), and hence does not affect the sum of the partners' payoffs.[15]

Before proceeding with a firm's problem of finding the optimal incentive weights, α and β, it is useful to comment on the partners' participation constraints. Recall that a partner will agree to play the game if his (expected) payoff given the compensation contracts is non-negative. It is easy to see that if the firm's economic profit (i.e., the sum of the partners' payoffs) is non-negative, then there always exists a lump sum transfer, η, between the firm's partners so that each partner gets at least zero payoff. Therefore, a sufficient condition for the partners' participation constraints to be satisfied is that the firm's economic profit is non-negative.

We first consider the firm's choice of revenue-sharing rules, α. Note that a partner's break-even price is $\underline{f}^t_H = \kappa^t(\beta^t_H)/\alpha^t_H$. Thus, to each value of β^t_H, there corresponds a revenue-sharing rule, $\alpha^t_H = \kappa^t(\beta^t_H)/\mathrm{K}^t(\beta^t_H)$, such that the partner's break-even price coincides with the firm's cost (namely $\underline{f}^t_H = \mathrm{K}^t_H$), and the partner's pricing strategy therefore also maximizes the firm's payoff. In this sense, such a revenue-sharing rule is optimal for the firm conditional on the given β.

We now consider the firm's choice of liability-sharing rules, β. Natural candidates are those that induce partners to perform audits efficiently. This leads to the following definition:

Definition 1: For a client of type $t \in \{C, S\}$, define $\beta^{t\circ}$ as the liability-sharing rule that maximizes the audit's value, denoted by $\Omega(\beta^t, t)$, which is measured as the sum of the shareholder surplus from purchasing the audit and the audit firm's economic profit from producing the audit. That is, $\beta^{t\circ}$ solves the following maximization problem:

$$\begin{aligned}\max_{\beta^t \in [0,\bar{\beta}]} \Omega(\beta^t, t) &\equiv (v^t(\beta^t) - f) + (f - \mathrm{K}^t(\beta^t)) \\ &= v^t(\beta^t) - \mathrm{K}^t(\beta^t) \\ &= (1 - \phi)\mu^t(e^{t*}(\beta^t), h^{t*}(\beta^t))I - e^{t*}(\beta^t) - h^{t*}(\beta^t), \qquad (14)\end{aligned}$$

---

[15] A bargaining game between the firm's partners will be called for to determine the lump sum transfer, η, which is not the focus of this paper.

where $\bar{\beta}$ is such that $e^{t*}(\beta) = \bar{e}$ and $h^{t*}(\beta) = 0$ for any $\beta \geq \bar{\beta}$ and $t \in \{C, S\}$.

Note in Equation (14) that the value of an audit, $\Omega(\beta^t, t)$, is equal to the expected savings due to the audit of wasteful investments in a bad company, minus the partners' joint effort costs in performing the audit. An efficient audit is one whose value is maximized by the liability-sharing rule β^{to} through the induced effort profile, and the value of an efficient audit equals $\Omega(e^{t*}(\beta^{to}), h^{t*}(\beta^{to}); t)$.

We assume enough regularity such that β^{to} is unique. In addition, assume that $\beta^{to} > 0$, which implies that the value of an efficient audit is positive, thus ensuring the existence of an audit market. Next, we assume that an efficient audit requires teamwork (namely $h > 0$). Based on Proposition 1, this assumption implies that $\beta^{to} < 1$.[16] That is, it is optimal for partners to share the legal liability costs of an audit in order to induce the partner collaboration required for the efficient performance of the audit. Finally, assume that the difference between the two types of clients—as characterized by the parameterized audit quality function $\mu(e, h; t)$—is nontrivial such that $\beta^{So} \neq \beta^{Co}$. This implies that the effort profile, (e, h), required to provide an efficient audit varies according to client type. In particular, using Topkis' monotonicity theorem and Equation (14), we observe that $\beta^{Co} < \beta^{So}$ when the following condition holds:

$$\frac{\partial\Omega(\beta, C)}{\partial\beta} < \frac{\partial\Omega(\beta, S)}{\partial\beta} \text{ for all } \beta \in (0, 1), \tag{15}$$

where $\partial\Omega(\beta,t)/\partial\beta = [(1 - \phi)\mu_e^t I - 1]e_\beta^{t*} + [(1 - \phi)\mu_h^t I - 1]h_\beta^{t*}$, $t \in \{C, S\}$, and the subscripts e, h, and β denote derivatives. Condition (15) states that the marginal benefit of β is greater for an S type client than for a C type client. Thus, when Condition (15) holds, a higher client type makes a larger β desirable; that is, the efficient audit of an S type client requires that greater legal liability costs be imposed on the engagement partner, compared with the efficient audit of a C type client.

The next proposition establishes the firms' equilibrium compensation strategies.

Proposition 3: Firms have the same dominant compensation strategy: For clients of type $t \in \{C, S\}$, the sharing rules are $(\alpha_{ri}^{t*}, \beta_{ri}^{t*}) = (\alpha^{to}, \beta^{to})$, where $r \in \{A, B\}$, $i \in \{1, 2\}$, $\alpha^{to} \equiv \kappa^t(\beta^{to})/K^t(\beta^{to})$, and κ^t and K^t are the engagement partner's personal cost and the firm's cost of an audit, respectively. In the dominant strategy equilibrium, firms share clients of each type and earn zero economic profit.

The equilibrium has the two firms setting the same compensation strategy, composed of client-type-specific sharing rules. The intuition for the equilibrium is as follows. The equilibrium revenue-sharing rule equates the partner's break-even price to the firm's cost for each type of client (i.e., $f^t = K^t$ when $\alpha^{to} \equiv \kappa^t/K^t$). As such, the revenue-sharing rule induces a partner to choose the audit fee to maximize the firm's payoff. Furthermore, under such a revenue-sharing rule, the maximum shareholder surplus for the partner equals the value of the audit (since $\sigma^t \equiv v^t - f^t = v^t - K^t \equiv \Omega^t$ when $f^t = K^t$); therefore, by maximizing the audit's value, the equilibrium liability-sharing rule, β^{to}, maximizes not only the partner's incentive to win an audit engagement, but also the firm's payoff from the audit. To see this, note from Proposition 2 that the firm's payoff from an audit it performs equals the audit's value less the rival auditor's maximum shareholder surplus (i.e., Firm r's payoff is $(v_r^t - \sigma_q^t) - K_r^t = \Omega_r^t - \sigma_q^t$, $r, q \in \{A, B\}$).

---

[16] Recall that Proposition 1 shows that an auditor is induced to provide cooperation to his partner if and only if he bears some responsibility for the legal liability costs of his partner's clients.

It is interesting to note that $\beta^{t\circ} \in (0, 1)$ implies that $\alpha^{t\circ} \in (0, 1)$. That is, given that it is optimal for partners to share the legal liability cost of an audit in order to induce teamwork for the audit, it is also optimal for partners to share the audit fee revenue of the audit. To explain revenue sharing, first note that $f^t = \kappa^t < K^t$ when $\alpha^t = 1$, which means that a partner tends to bid below the firm's cost (i.e., $f^t < K^t$) when he gets the entire fee revenue from his clients (i.e., when $\alpha^t = 1$). Next, note that $f^t = \kappa^t/\alpha^t = K^t$ when $\alpha^t = \kappa^t/K^t < 1$, which implies that sharing revenues (i.e., $\alpha^t < 1$) has an effect of imposing "taxes" on engagement partners in order to counteract the individual partner's inappropriate aggressiveness with regard to his pricing policy, thereby aligning the partner's break-even price, f^t, with the firm's cost, K^t.

We conclude this section by noting that the equilibrium is a perfect competition equilibrium: the same compensation schemes by the firms induce the same effort and pricing strategies. As a result, for each client, auditors from different firms bid the same price (which equals the firm's cost), have an equal chance to win the bid, and earn zero economic profit. In equilibrium, the two firms have the same market share in each location, namely $(n_i^C + n_i^S)/2$, $i \in \{1, 2\}$. Shareholders benefit the most from the perfect competition in that they collect the entire value of the audits.

IV. CLIENT TYPE NOT CONTRACTABLE

The perfect competition result in the previous section hinges on the assumption that client type is contractable, which as discussed in Section II, we believe, is implausible in the real world. In this section we examine the case in which client type is not contractable. Specifically, we assume that a client's type is observable to the parties who are actively involved in the audit (including representatives of the shareholders who hire auditors, the engagement partner, and the partner who provides assistance in the performance of the audit), but cannot be verified by an enforcement party such as the head office of the audit firm or a court, and thus cannot be used for partner compensation purposes.[17] Accordingly, the firm's Compensation Contract (9) cannot vary with client type and becomes:

$$y_i = \alpha_i \sum_{t\in\{C,S\}} R_i^t + (1-\alpha_j) \sum_{t\in\{C,S\}} R_j^t - \beta_i \sum_{t\in\{C,S\}} L_i^t - (1-\beta_j) \sum_{t\in\{C,S\}} L_j^t + \eta_{ij}, \tag{16}$$

where $i \neq j$, $i, j \in \{1, 2\}$.

Analysis and Results

Since the effort choice and the bidding subgames are the same as before, the effort and pricing strategies—as functions of the profit-sharing rules—are the same as characterized in Propositions 1 and 2. The critical difference between this and the previous case lies in the contracting stage. In the previous case, a firm can use client-type-specific sharing rules to induce optimal partner behavior for all types of clients; but when client type is not contractable, the firm cannot find uniform sharing rules that induce the same optimal behavior. To see this, note that a liability-sharing rule, β, does not exist that induces efficient audits for all clients because $\beta^{C\circ} \neq \beta^{S\circ}$. Despite this apparent contracting disadvantage, careful examination of interactions between firms reveals that the firms may differentiate themselves by specializing in different types of clients, thereby gaining limited monopoly power over the clients in which they specialize. Proposition 4 identifies conditions under which there exists a specializing equilibrium defined as follows:

---

[17] Partners must be able to *verify* the client's type in order to implement a type-dependent sharing rule. Although the client's type is observable to the engagement partner and the helping partner, both of whom are actively involved in the audit and hence have learned the client's type in the course of doing business, they have incentives to misreport the client type *ex post* (e.g., a helping partner tends to claim a simple client as a complex client to seek more compensation for his cooperation effort). Verifiability implies that an independent third party, such as a court or other partners in a multiple-partner firm, can easily settle compensation-related disputes.

Definition 2: In Location $i \in \{1, 2\}$, a specializing equilibrium is characterized by the firms' liability-sharing rules $(\beta_{qi}, \beta_{ri}) = (\beta^{S\circ}, \beta^{C\circ})$, $q \neq r$, $q, r \in \{A, B\}$. That is, in equilibrium, firms specialize in different types of clients by choosing different value-maximizing, liability-sharing rules. The type of clients in which a firm specializes constitutes the firm's niche market.

Proposition 4: There exists a specializing equilibrium where $\beta^*_{Ai} = \beta^*_A = \beta^{S\circ}$, $\beta^*_{Bi} = \beta^*_B = \beta^{C\circ}$, $\alpha^*_{Ai} = \alpha^*_A = \alpha^{S\circ}$, $\alpha^*_{Bi} = \alpha^*_B = \alpha^{C\circ}$, $i \in \{1, 2\}$, if the following conditions hold:

$$v^S(\beta^{S\circ}) - K^S(\beta^{S\circ}) > v^S(\beta^{C\circ}) - \frac{\kappa^S(\beta^{C\circ})K^C(\beta^{C\circ})}{\kappa^C(\beta^{C\circ})}, \tag{17}$$

$$v^C(\beta^{C\circ}) - K^C(\beta^{C\circ}) > v^C(\beta^{S\circ}) - \frac{\kappa^C(\beta^{S\circ})K^S(\beta^{S\circ})}{\kappa^S(\beta^{S\circ})}, \tag{18}$$

In equilibrium, each firm audits only clients from its niche market and earns positive economic profits as follows:

$$\prod\nolimits_A = (\min\{v_A^{S*}, v_A^{S*} - \sigma_B^{S*}\} - K_A^{S*})(n_1^S + n_2^S), \tag{19}$$

$$\prod\nolimits_B = (\min\{v_B^{C*}, v_B^{C*} - \sigma_A^{C*}\} - K_B^{C*})(n_1^C + n_2^C), \tag{20}$$

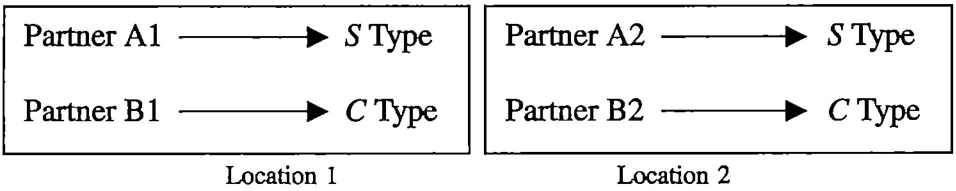
where Π denotes the firm's economic profit, and n_i^t is the number of clients of type $t \in \{C, S\}$ in Location $i \in \{1, 2\}$.

Panel A of Figure 3 depicts the equilibrium established in Proposition 4, where the firms specialize in different types of clients: Firm A's market share is $n_1^S + n_2^S$ and Firm B's market share is $n_1^C + n_2^C$. In equilibrium, Firm A (Firm B) chooses the liability-sharing rule that maximizes the value of audits for the S type (C type) clients, so that auditors from Firm A (Firm B) become specialists for the S type (C type) clients. Note that an auditor acting as a specialist for a type of client does not mean that he has specialized skills at performing the audit, but rather that—because of his firm's compensation schemes—the auditor and his partner are induced to provide a more efficient audit (i.e., a higher value audit) for that type of client than would the auditors from the rival firm.

We now explain the sufficient conditions, Inequalities (17) and (18), under which a specializing equilibrium exists. Note that the specializing liability-sharing rules (that is, efficient audits) alone do not ensure that auditors win the clients in which they specialize.

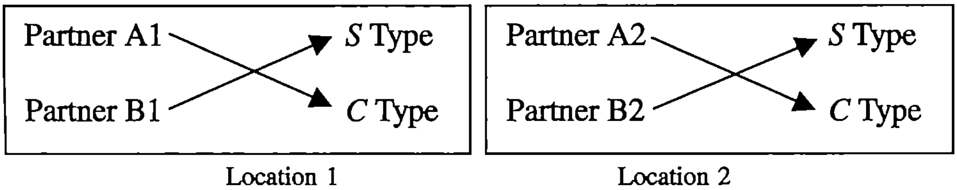
FIGURE 3
Specializing Equilibria

Panel A: Partners within a Firm are Compensated in the Same Way



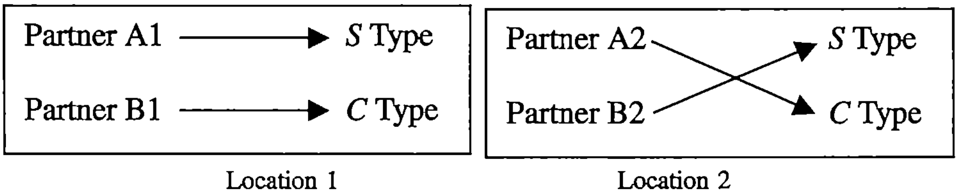
For the numerical example given in the text, in this case, $\alpha^*_{A1} = \alpha^*_{A2} = 0.9994$, $\beta^*_{A1} = \beta^*_{A2} = 0.998$, $\alpha^*_{B1} = \alpha^*_{B2} = 0.8807$, $\beta^*_{B1} = \beta^*_{B2} = 0.8677$. Fix $n^t_i = 10$ where $i \in \{1, 2\}$ and $t \in \{C, S\}$; the total shareholder surplus is 17.13 and Firm A's (Firm B's) economic profits are 1.584 (0.068).

Panel B: Partners within a Firm are Compensated in the Same Way



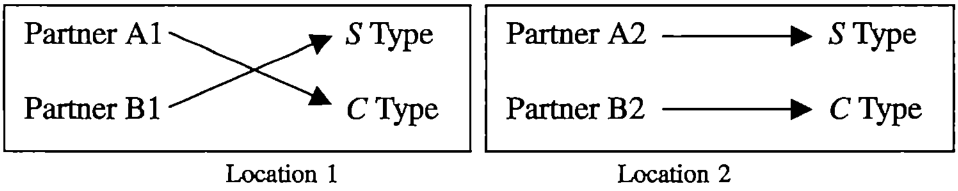
For the numerical example given in the text, in this case, $\alpha^*_{A1} = \alpha^*_{A2} = 0.8807$, $\beta^*_{A1} = \beta^*_{A2} = 0.8677$, $\alpha^*_{B1} = \alpha^*_{B2} = 0.9994$, $\beta^*_{B1} = \beta^*_{B2} = 0.998$. Fix $n^t_i = 10$ the total shareholder surplus is 17.13 and Firm A's (Firm B's) economic profits are 0.068 (1.584).

Panel C: Partners within a Firm are Compensated in Different Ways



For the numerical example given in the text, in this case, $\alpha^*_{A1} = \alpha^*_{B2} = 0.9994$, $\beta^*_{A1} = \beta^*_{B2} = 0.998$, $\alpha^*_{B1} = \alpha^*_{A2} = 0.8807$, $\beta^*_{B1} = \beta^*_{A2} = 0.8677$. Fix $n^t_i = 10$ the total shareholder surplus is 17.13 and Firm A's (Firm B's) economic profits are 0.826 (0.826).

Panel D: Partners within a Firm are Compensated in Different Ways



For the numerical example given in the text, in this case, $\alpha^*_{A1} = \alpha^*_{B2} = 0.8807$, $\beta^*_{A1} = \beta^*_{B2} = 0.8677$, $\alpha^*_{B1} = \alpha^*_{A2} = 0.9994$, $\beta^*_{B1} = \beta^*_{A2} = 0.998$. Fix $n^t_i = 10$; the total shareholder surplus is 17.13 and Firm A's (Firm B's) economic profits are 0.826 (0.826).

We know from Proposition 2 that the bidding outcome depends on the maximum shareholder surplus, which is based not only on the induced effort profile, but also on the induced pricing policy. More specifically, although the revenue-sharing rule of Firm A induces the optimal pricing strategy for the *S* type clients, it cannot do so for the *C* type clients at the same time since $\alpha^{S\circ} \equiv \kappa^S(\beta^{S\circ})/K^S(\beta^{S\circ}) \neq \kappa^C(\beta^{S\circ})/K^C(\beta^{S\circ})$ (which implies that the break-even price differs from the firm's cost for the *C* type clients, i.e., $\underline{f}_A^C \neq K_A^C$). Thus, Firm A must be concerned that its partners would attract the *C* type clients with prices below the firm's cost. This situation would be prevented by Condition (18), which states that the maximum shareholder surplus of Firm A for a *C* type client (the right-hand side) is less than the maximum shareholder surplus of Firm B for a *C* type client (the left-hand side, which equals the value of an efficient audit), i.e., $\sigma_A^{C*} < \sigma_B^{C*}$ (see the Appendix for details). In other words, Firm A's sharing rules are such that its partners—who are nonspecialists for a *C* type client—would not be able to provide a shareholder surplus greater than the value of an efficient audit provided by Firm B for a *C* type client. It is easy to show that if Condition (18) does not hold, then Firm A's partners would bid below the firm's cost for a *C* type client; that is, if $\sigma_B^C \leq \sigma_A^C$, then $f_A^C < K_A^C$. It follows that Condition (18) holds when Firm A's partners are induced not to bid below the firm's cost for Firm B's clientele. A similar analysis applies to Firm B and Condition (17). In sum, these conditions impose restrictions on the parameters of the model such that partners are induced not to bid far below their firm's cost for the rival firm's clientele.[18] It should be noted that Conditions (17) and (18) are expressed in terms of β'°, which itself is a function of the primitive parameters of the model (see Definition 1). While the complexity of the multistage problem prohibits expressing the sufficient conditions in terms of the primitive parameters of the model, we later use a numerical example to illustrate cases in which these conditions hold.

We observe that neither firm has an incentive to deviate from the equilibrium in which the firms obtain positive payoffs (namely economic profits) from specialization. To explain this, suppose that Firm A wants to appropriate Firm B's payoff. Firm A must first change its liability-sharing rule so as to provide higher value audits for Firm B's clientele. Since Firm B has been efficient (and its partners are induced to bid in the firm's interest) with respect to its clientele, Firm A cannot earn any economic profit from the rival firm's niche market. However, changing its liability-sharing rule will cause Firm A to provide lower value audits for its own niche market and to earn less economic profit from each of its clients, which equals the decreased value of the audit less the rival auditor's maximum shareholder surplus (if $\sigma_A^S \geq \sigma_B^{S*}$) or equals 0 (if $\sigma_A^S < \sigma_B^{S*}$). A similar analysis applies to Firm B. As a result, positive equilibrium payoffs will not be bid away in competition.

Although the total value of audits in the specializing equilibrium is the same as in the perfect competition equilibrium analyzed in Section III, the audit's value with specialization is divided between the audit firms and the shareholders; in contrast, all the value of audits accrues to the shareholders under perfect competition.

Example

To help understand the equilibrium results, consider the following example. Suppose the audit quality function is $\mu(e, h; x) = x\sqrt{e} + (1 - x)\sqrt{eh}$, where $e \in [0, 1]$, $h \in [0, 1]$, $x \in (0, 1)$. In addition, take $C = 0.5$, $S = 0.95$, $\phi = 0.7$, $l = I = 6$. Thus, we have $\mu^C(e, h) = 0.5\sqrt{e} + 0.5\sqrt{eh}$ and $\mu^S(e, h) = 0.95\sqrt{e} + 0.05\sqrt{eh}$.

---

[18] That partners are induced not to bid below the firm's cost for the rival firm's clientele is a more stringent condition than Conditions (17) and (18). In fact, Conditions (17) and (18) hold even when partners bid below the firm's cost for the rival firm's clientele (e.g., $\underline{f}_A^C < K_A^C$), as long as the bid is not *far below* the firm's cost so that the resulting shareholder surplus does not exceed the value of an efficient audit provided by the rival firm.

It can be shown that when client type is contractable, the dominant strategy equilibrium composed of client-type-specific sharing rules is as follows: $\{\alpha_{ri}^{S*} = 0.9994, \beta_{ri}^{S*} = 0.998, \alpha_{ri}^{C*} = 0.8807, \beta_{ri}^{C*} = 0.8677\}$ for all $r \in \{\text{A, B}\}$ and $i \in \{1, 2\}$. In contrast, when client type is not contractable, there exists a specializing equilibrium as follows (see Panel A of Figure 3): $\{\alpha_{Ai}^{*} = 0.9994, \beta_{Ai}^{*} = 0.998; \alpha_{Bi}^{*} = 0.8807, \beta_{Bi}^{*} = 0.8677\}$, and the induced effort choices are $\{e_{Ai}^{S*} = 0.7281, h_{Ai}^{S*} = 0.5898 \times 10^{-8}; e_{Bi}^{C*} = 0.1598, h_{Bi}^{C*} = 0.566 \times 10^{-3}\}$ for all $i \in \{1, 2\}$.

As shown in Panel A of Figure 3, in the specializing equilibrium, Firm A (Firm B) audits only the simple (complex) clients. Observe that $\alpha_{Ai}^{*} > \alpha_{Bi}^{*}$ and $\beta_{Ai}^{*} > \beta_{Bi}^{*}$, $i \in \{1, 2\}$, which means that Firm A, which specializes in the simple clients, bases a partner's compensation more on his local office performance—as reflected by the revenues and liabilities associated with the partner's own audit engagements—than does Firm B, which specializes in the complex clients. We first explain why $\beta_A^* > \beta_B^*$. Note that Condition (15) is satisfied, which as indicated earlier, implies that a larger (smaller) β is desirable for the efficient audit of a simple (complex) client. The factors that work together for Condition (15) to hold are rather complicated. We here offer a discussion at the most intuitive level by focusing only on the dominant effects. As implied by Assumption 3, the marginal benefit of cooperation effort h (engagement audit effort e) is greater for the complex client (for the simple client). Intuitively, this suggests that the efficient audit of a complex client requires more cooperation/helping effort than that of a simple client. Consistent with this intuition, Firm B imposes a larger share of the legal liability costs on helping partners (namely a smaller β) to induce greater helping effort, compared with Firm A (i.e., $h_B^{C*} > h_A^{S*}$).

We next explain why $\alpha_A^* > \alpha_B^*$. It is straightforward to verify that $\kappa_A^{S*}/K_A^{S*} > \kappa_B^{C*}/K_B^{C*}$, which means that there is a relatively larger discrepancy between the engagement partner's personal cost and the firm's cost of an audit in Firm B than in Firm A with respect to their clientele; therefore, a partner of Firm B would be relatively more aggressive in the pricing policy if he were fully compensated for each dollar of fee revenues he brought into the firm (i.e., if $\alpha = 1$). It follows that Firm B imposes a higher "tax rate" on engagement partners (namely a smaller α) to induce the optimal pricing behavior for complex clients than Firm A does for simple clients. In essence, this example illustrates that, when the client's type is not contractable, each firm uses type-oriented sharing rules to induce the partner behavior (in terms of the partner collaboration and pricing policy) most suitable for the type of clients in which the firm specializes.

Given that $\phi = 0.7$ and $l = I = 6$, while we cannot check the sufficient Conditions (17) and (18) for all the possible values of $\{C, S\}$, where $0 < C < S < 1$, we can identify a large set of values of $\{C, S\}$ for which the sufficient conditions hold and thus a specializing equilibrium exists. Examples include the following: (i) Fix $S = 0.95$; the sufficient conditions hold for all $C \in \{0.14, 0.15, \ldots, 0.73, 0.74\}$. (ii) Fix $S = 0.99$ the sufficient conditions hold for all $C \in \{0.03, 0.04, \ldots, 0.9, 0.91\}$. (iii) The sufficient conditions hold for all $(C, S) \in \{(C, S) | C \in \{0.26, 0.27, \ldots, 0.59, 0.60\}, S \in \{0.92, 0.93, \ldots, 0.98, 0.99\}\}$.[19]

Other Issues

In this subsection we discuss two technical issues and limitations related to our equilibrium results: the multiplicity of specializing equilibria, and the observability of compensation contracts. First, we note that, under the conditions specified in Proposition 4, there exist four specializing equilibria as shown in Figure 3. In addition to the two equilibria in which partners within a firm are compensated in the same way—as depicted in Panels A and B of Figure 3—the other two equilibria are such that partners within a firm are compensated differently so that each office of the firm specializes in a different type of client, as depicted in Panels C and D of Figure 3. For example, the following is also an equilibrium (see Panel C):

---

[19] Our study of many hundreds of numerical examples suggests that the sufficient conditions are more likely to hold when the client types, C and S, are far apart.

$$\{\alpha^*_{A1} = \alpha^*_{B2} = \alpha^{So}, \beta^*_{A1} = \beta^*_{B2} = \beta^{So}, \alpha^*_{B1} = \alpha^*_{A2} = \alpha^{Co}, \beta^*_{B1} = \beta^*_{A2} = \beta^{Co}\}.$$

That is, in Location 1, Partner A1 (Partner B1) specializes in S type (C type) clients, whereas in Location 2, Partner A2 (Partner B2) specializes in C type (S type) clients.

It is straightforward to see that the economic profit earned from each client is the same across the four specializing equilibria, regardless of which firm earns the profit. Therefore, the sum of the two firms' economic profits is the same across the equilibria. To illustrate, we return to the preceding numerical example. It can be shown that in the dominant equilibrium analyzed in Section III, for $n_i^t = 10$ where $i \in \{1, 2\}$ and $t \in \{C, S\}$, the total shareholder surplus is 18.782 and each audit firm earns zero economic profit. By comparison, when client type is not contractable, for $n_i^t = 10$, Figure 3 demonstrates the total shareholder surplus and the firms' economic profits for each of the four specializing equilibria; clearly, moving from one to another equilibrium only makes one firm better off and the other worse off, with the total of the firms' payoffs unchanged. It can be shown that these are undominated equilibria (i.e., firms play undominated strategies in equilibrium).[20] In this case, game theory does not offer a general theory about which of the equilibria will be obtained.

In these specializing equilibria, firms achieve monopoly positions for the clients in which they specialize. However, it should be noted that in the real world, there may exist mismatches in the sense that a small proportion of an audit firm's clients may come from a rival firm's niche market. Admittedly, our model does not consider all the factors that affect a client's auditor choice. For example, client companies may desire audit firms not associated with their rival companies for fear of revealing proprietary information to their competitors (Kwon 1996). In addition, client companies may not want to hire their executives' former CPA firms as auditors for fear that audit quality will be impaired (Lennox and Park 2003). Moreover, not all companies' auditor choices are determined by independent, competent audit committees (i.e., the auditor choices are not always optimal), which can also cause real-world mismatches.

The second issue we consider is observability of contracts. A crucial assumption underlying the specializing equilibria is that firms' compensation plans are observable to all game players. This may occur, for example, by learning through repeated play of the game. Although this assumption is commonplace in the economics literature on strategic delegation (e.g., the Fershtman and Judd [1987] and Sklivas [1987] type models), it is often criticized on the grounds that the contracts may not be renegotiation-proof (see Katz 1991). That is, the equilibrium may not hold if a firm secretly changes its own contracts after all firms have publicly announced their contracts. Unfortunately, we cannot solve for an equilibrium in which client type is not contractable and compensation contracts are not publicly observable.[21] In this regard, we note that our study differs from Koçkesen and Ok (2004), who find that, when the decision of delegating or not delegating is publicly observable, there may exist a delegation equilibrium in which unobservable compensation contracts can play a strategic role. While the delegation of a job by a principal to an agent is an option in their study, decentralization is a must in our study since the "principal" here is not a person, but rather represents the collective interest of a partnership in which individual partners make decentralized effort decisions.

---

[20] The proof that these equilibria are undominated is essentially the same as the proof of Liu (2002, Corollary 3.1) and is thus omitted.

[21] However, it can be shown that the specializing equilibria that exist under the conditions specified in Proposition 4 cease to hold when compensation contracts are unobservable.

V. CONCLUDING REMARKS

In this paper, we examine the effects of compensation arrangements on the behavior of partners within a public accounting firm, as well as on the strategic interactions of the firms in an auditing oligopoly. We use a stylized model where (among other assumptions) there are two audit firms and two types of clients. We show that when partner profit-sharing rules can vary with client type (i.e., client type is contractable), a perfect competition equilibrium will obtain. However, when the client's type is not contractable—which we believe to be a reasonable description of the real world—firms may strategically differentiate themselves by choosing different profit-sharing rules in order to specialize in different types of clients. By doing so, the firms avoid cutthroat competition and earn positive economic profits from their niche markets. Our equilibrium results generalize naturally to the situation in which there are m audit firms and m different types of clients (where $m > 2$). The idea can also be applied to the analysis of the scenario where there are many client types in an auditing oligopoly.

The theory developed in this paper offers an explanation for the observed differences in the compensation plans of the big public accounting firms; namely, there exist both "small profit pool" and "large profit pool" firms. In the former, local office profitability is very important to the compensation of a partner, while local office profitability is not important in the latter case. An implication of the situation where client type is not contractable and firms adopt different compensation arrangements in a specializing equilibrium is that the client portfolios of firms will vary. In the year 2000, we conducted structured interviews with partners from the Vancouver, Canada offices of KPMG and (former) Arthur Andersen & Co. to investigate the compensation arrangements of those two big firms. Based on these interviews, we can classify KPMG (Arthur Andersen) as a "small pool" ("large pool") firm in the sense of Trompeter (1988). Thus, KPMG's sharing rules are more tailored to the audits of simple companies, which do not require much partner cooperation, whereas Arthur Andersen's sharing rules are more suitable for the audits of relatively large complex companies, whose audits require more partner cooperation. These partner compensation arrangements can, for example, explain the U.S. Senate survey results (U.S. Senate 1977, 30–31, 421) regarding the client portfolios of the then Big 8 firms. In that survey, Arthur Andersen (Peat Marwick International, now KPMG) reported that, as of 1975, it had 417 (331) audit clients that were corporations listed on the NYSE or AMSE, in contrast to 6,000 (23,245) audit clients that were either not listed on the major stock exchanges or were privately held corporations. This evidence suggests that, consistent with the implications of their compensation plans, Arthur Andersen was relatively specialized in the large/complex company sector, and Peat Marwick International (now KPMG) was relatively specialized in the small/simple company sector.[22] We also note that 19 years later, O'Keefe et al. (1994) found that, in an industry-stratified random sample of U.S. audit clients of KPMG, fully 80 percent were closely held companies.

---

[22] Similar observations are made by Liu (2002) based on *Who Audits America*.

To conclude, while the above evidence is suggestive and consistent with the theory developed in this paper, future empirical studies would be useful to examine the details of the actual profit-sharing arrangements among the partners in each of the now Big 4 firms. Differences in compensation arrangements can then be studied to determine their impact on audit firm decisions and firm characteristics (through the effects on individual partners' behavior), such as their client portfolios and audit quality. Such studies will help us gain a greater understanding of the functioning of audit firms in the economy.

APPENDIX

See Table 1 for notation definitions.

Proof of Proposition 1

In the audit firm, consider the case $\beta_i^t \in (0, 1)$, $i \in \{1, 2\}$. It is easy to see from Assumptions 1, 2, 4, and 6 that, for any of Partner *i*'s clients of type $t \in \{C, S\}$, the reaction curves (i.e., partners' effort choice functions), $e(h)$ and $h(e)$, are determined by (A1) and (A2), respectively:

$$\beta_i^t(1 - \phi)\mu_e^t(e, h)l - 1 = 0, \tag{A1}$$

$$(1 - \beta_i^t)(1 - \phi)\mu_h^t(e, h)l - 1 \leq 0, \text{ with equality if } e > 0. \tag{A2}$$

We first prove the existence of equilibrium. Denote by $\hat{h}$ the level of effort by the helping partner given that the engagement partner's effort level attains the upper bound, i.e., $\hat{h} = h\ (\bar{e})$, which is less than $\bar{h}$ by Assumption 6. Let $h^{-1}(\cdot)$ represent the inverse function of $h(\cdot)$. Note that $e(h)$ and $h^{-1}(h)$ are continuous and strictly increasing on $[0, \hat{h}]$ by Assumptions 4 and 5. Consider the function ζ defined by $\zeta(h) = e(h) - h^{-1}(h)$.. $\zeta(0) > 0$ because $e(0) > 0$ and $h^{-1}(0) = 0$. On the other hand, $\zeta(\hat{h}) < 0$ because $e(\hat{h}) < \bar{e}$ and $h^{-1}(\hat{h}) = \bar{e}$. Since ζ is continuous on $[0, \hat{h}]$, according to the intermediate value theorem there exists $h' \in (0, \hat{h})$ such that $\zeta(h') = 0$, namely $e(h') = h^{-1}(h')$. Thus, the reaction curves intersect in $(0, \hat{h})$. This proves the existence of equilibrium, and that $e_i^{t*} \in (0, \bar{e})$ and $h_j^{t*} \in (0, \bar{h})$, where $i \neq j$, $i, j \in \{1, 2\}$.

We now prove the uniqueness of the equilibrium. A sufficient condition for the solution to be unique (i.e., for the reaction curves to intersect only once) is that the reaction curve of the engagement partner is flatter than that of the helping partner at equilibrium. This condition is satisfied given Assumptions 4 and 5. Assumption 4 states that $\mu_{ee}\mu_{hh} - \mu_{eh}\ \mu_{eh} > 0$. Thus, we know from Assumption 5 that $-\mu_{eh}/\mu_{ee} < -\mu_{hh}/\mu_{eh}$, which then implies $\partial e(h)/\partial h < \partial h^{-1}(h)/\partial h$.

Therefore, when $\beta_i^t \in (0, 1)$, there exists a unique equilibrium (e_i^{t*}, h_j^{t*}) jointly determined by Equations (12) and (13). The proof of the equilibrium in other cases (i.e., $\beta_i^t = 0$ and $\beta_i^t \geq 1$) is straightforward and hence omitted.

Proof of Proposition 2

Consider the local competition for a client of type $t \in \{C, S\}$ in Location $i \in \{1, 2\}$. Shareholders will not purchase an audit if they get a negative surplus. Therefore, if $\sigma_{qi}^t < 0$ and $\sigma_{ri}^t < 0$, where $q \neq r$, $q, r \in \{\text{A, B}\}$, then both auditors have zero demand. If $\sigma_{qi}^t < 0$ and $\sigma_{ri}^t \geq 0$, then Auditor *qi* has zero demand, and Auditor *ri* becomes a monopolist, who charges the price of υ_{ri}^t to provide just zero shareholder surplus. In this monopoly case, Auditor *ri*'s payoff is $\alpha_{ri}^t\upsilon_{ri}^t - \kappa_{ri}^t \geq 0$, where the inequality (equality) follows from $\sigma_{ri}^t = \upsilon_{ri}^t - \kappa_{ri}^t/\alpha_{ri}^t \geq 0$.

If $\sigma^t_{qi} = \sigma^t_{ri} \geq 0$, then it is easy to verify that the only equilibrium is for both auditors to bid their break-even prices. Since shareholders are indifferent between the two bids, it follows from Assumption 7 that the auditors win the bid with equal chance.

If $0 \leq \sigma^t_{qi} < \sigma^t_{ri}$, then we propose that the strategy profile $(f^{t*}_{qi}, f^{t*}_{ri}) = (\underline{f}^t_{qi}, v^t_{ri} - \sigma^t_{qi})$ is a unique (pure strategy) equilibrium. Note that in equilibrium, shareholders are indifferent between the two bids. Thus, Auditor *ri* wins the bid by Assumption 7. In this case, Auditor *ri*'s payoff is $\alpha^t_{ri}(v^t_{ri} - \sigma^t_{qi}) - \kappa^t_{ri} > 0$, where the inequality follows from $\sigma^t_{ri} = v^t_{ri} - \kappa^t_{ri}/\alpha^t_{ri} > \sigma^t_{qi}$. We now show that the above strategy profile is an equilibrium by showing that an auditor cannot gain from a unilateral deviation from the equilibrium. Given Auditor *ri*'s equilibrium strategy, Auditor *qi* would still have zero demand if he increased the bid, and would suffer a loss if he decreased the bid. Given Auditor *qi*'s equilibrium strategy, Auditor *ri* would lose the profitable bid if he increased the bid, and would leave money on the table if he decreased the bid. Furthermore, the (pure strategy) equilibrium is unique given that an auditor would not bid below his break-even price. To see this, suppose that Auditor *qi* bids above his break-even price: $f^{t'}_{qi} > \underline{f}^t_{qi}$. This cannot be an equilibrium strategy: Auditor *ri*'s best response is to bid a price at $v^t_{ri} - (v^t_{qi} - f^{t'}_{qi})$, but Auditor *qi* could then increase his payoff by choosing a price between his break-even price and $f^{t'}_{qi}$.

Proof of Proposition 3

Due to limited space, we only lay out the main steps of the proof. Interested readers can refer to Liu (2002) for the details.

For clients of type $t \in \{C, S\}$, consider Firm A's problem with respect to Partner A1. The proof that $(\kappa^t(\beta^{to})/K^t(\beta^{to}), \beta^{to})$ is the dominant strategy for Firm A with respect to Partner A1 takes the following three steps:

The first step is to show that $(\alpha^t(\beta^t_{A1}), \beta^t_{A1}) \equiv (\kappa^t(\beta^t_{A1})/K^t(\beta^t_{A1}), \beta^t_{A1})$ dominates $(\alpha^{t'}_{A1}, \beta^t_{A1}$ for all $\alpha^{t'}_{A1} \neq \alpha^t(\beta^t_{A1})$ and $\beta^t_{A1} \in \{\beta | v^t(\beta) - K^t(\beta) > 0\}$.

The second step is to show that $(\kappa^t(\beta^{to})/K^t(\beta^{to}), \beta^{to})$ dominates $(\alpha^t(\beta^t_{A1}), \beta^t_{A1})$ for all $\beta^t_{A1} \in \{\beta | \beta \neq \beta^{to} \text{ and } v^t(\beta) - K^t(\beta) > 0\}$.

The third step is to show that $(\kappa^t(\beta^{to})/K^t(\beta^{to}), \beta^{to})$ dominates $(\alpha^t_{A1}, \beta^t_{A1})$ for all α^t_{A1} and $\beta^t_{A1} \in \{\beta | v^t(\beta) - K^t(\beta) \leq 0\}$.

Similarly, we can show that $(\kappa^t(\beta^{to})/K^t(\beta^{to}), \beta^{to})$ is the dominant strategy for Firm A with respect to Partner A2. The proof for Firm B is exactly the same. A combination of dominant strategies is necessarily a Nash equilibrium. In equilibrium, the maximum shareholder surplus provided by the rival auditors is the same for each client; therefore, based on Proposition 2, the auditors bid their break-even prices and share clients of each type. Note that under the equilibrium revenue-sharing rule $\alpha^{to} \equiv \kappa^t(\beta^{to})/K^t(\beta^{to})$, the break-even price equals the firm's cost, i.e., $f^{t*} = K^t(\beta^{to})$. It follows that in equilibrium each firm earns zero economic profit, and thus the partners' participation constraints are satisfied.

Proof of Proposition 4

Given the equilibrium sharing rules $\beta^*_{Ai} = \beta^*_A = \beta^{So}$, $\beta^*_{Bi} = \beta^*_B = \beta^{Co}$, $\alpha^*_{Ai} = \alpha^*_A = \alpha^{So}$, $\alpha^*_{Bi} = \alpha^*_B = \alpha^{Co}$, $i \in \{1, 2\}$, Conditions (17) and (18) imply that:

$$v^{S*}_A - K^{S*}_A > v^{S*}_B - \kappa^{S*}_B K^{C*}_B / \kappa^{C*}_B \text{ and } v^{C*}_B - K^{C*}_B > v^{C*}_A - \kappa^{C*}_A K^{S*}_A / \kappa^{S*}_A. \tag{A3}$$

Since (A3) holds for all partners within the firm, we have dropped the subscript *i*. Also note that $\sigma^{S*}_A = v^{S*}_A - K^{S*}_A > 0$, $\sigma^{S*}_B = v^{S*}_B - \kappa^{S*}_B K^{C*}_B / \kappa^{C*}_B$, $\sigma^{C*}_B = v^{C*}_B - K^{C*}_B > 0$, and $\sigma^{C*}_A = v^{C*}_A - \kappa^{C*}_A K^{S*}_A / \kappa^{S*}_A$. Thus, (A3) can be restated as (A4) or (A5):

$$\sigma_A^{S*} > \sigma_B^{S*} \text{ and } \sigma_B^{C*} > \sigma_A^{C*}. \tag{A4}$$

$$\upsilon_A^{S*} - \sigma_B^{S*} > \mathrm{K}_A^{S*} \text{ and } \upsilon_B^{C*} - \sigma_A^{C*} > \mathrm{K}_B^{C*}. \tag{A5}$$

(A4) and Proposition 2 imply that Firm A (Firm B) audits all the *S* type (*C* type) clients. Moreover, (A5) and Proposition 2 imply that each firm earns positive economic profits from its niche market as shown by Equations (19) and (20), and thus the partners' participation constraints are satisfied.

We now show that such sharing rules constitute an equilibrium. Observe that under these sharing rules, the value of audits is maximized for all auditors with respect to their clientele, and all auditors are induced to bid in their firms' interests with respect to their clientele. Then, it is straightforward to verify that, given the rival firm's equilibrium sharing rules, Firm A (Firm B) cannot earn any economic profits from the *C* type (*S* type) clients or increase its economic profits from the *S* type (*C* type) clients by unilaterally changing its own sharing rules for any partner. Hence, Conditions (17) and (18) are sufficient conditions for these sharing rules to form an equilibrium.

TABLE 1
Table of Notation

ϕ = probability of a company being a good company;
V = current liquidating value of a good company;
V' = end-of-period value of a good company;
I = current liquidating value of a bad company;
e = audit effort by engagement auditor;
h = cooperation/helping effort by an auditor for his partner's audit engagement;
g = auditor's report that the company is good;
b = auditor's report that the company is bad;
μ = audit quality, i.e., the (conditional) probability of detecting a bad company;
C = complex type of client;
S = simple type of client;
t = client type, $t \in \{C, S\}$;
x = parameter of the audit quality function, $\mu(\cdot)$, $x \in (\underline{x}, \bar{x})$, $\{C, S\} \in (\underline{x}, \bar{x})$;
l = legal payment made by an audit firm to shareholders in case of litigation;
f = audit fee;
$\underline{f}$ = auditor's break-even price for an audit engagement;
y = auditor's compensation;
R = audit fee revenues associated with an auditor's audit engagements;
L = legal liability costs associated with an auditor's audit engagements;
α = auditor's share of fee revenues from his own audit engagements;
β = auditor's share of liability costs from his own audit engagements;
η = lump sum transfer between the partners of a firm;
κ = engagement auditor's personal cost of an audit;
K = firm's cost of an audit;
υ = gross benefit of an audit to shareholders;
σ = maximum amount of shareholder surplus an auditor can provide;
Ω = value of an audit;
Π_r = economic profit of Firm $r \in \{\mathrm{A}, \mathrm{B}\}$; and
n_i^t = number of t type clients in Location $i \in \{1, 2\}$, $n_i^C + n_i^S = n_i$.

REFERENCES

Aggarwal, R., and A. Samwick. 1999. Executive compensation, strategic competition, and relative performance evaluation. *Journal of Finance* 54 (6): 1999–2043.

Burrows, G., and C. Black. 1998. Profit sharing in Australian Big 6 accounting firms: An exploratory study. *Accounting, Organizations, and Society* 23 (5-6): 517–530.

Carcello, J., D. Hermanson, and H. Huss. 2000. Going-concern opinions: The effects of partner compensation plans and client size. *Auditing: A Journal of Practice & Theory* 19 (1): 67–77.

Cypert, S. A. 1991. *Following the Money: The Inside Story of Accounting's First Mega-Merger.* New York, NY: AMACOM.

Dye, R. A. 1995. Incorporation and the audit market. *Journal of Accounting and Economics* 19: 75–114.

Ferguson, A., J. Francis, and D. Stokes. 2003. The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting Review* 78 (2): 429–448.

Fershtman, C., and K. Judd. 1987. Equilibrium incentives in oligopoly. *American Economic Review* 77 (5): 927–940.

———, ———, and E. Kalai. 1991. Observable contracts, strategic delegation, and cooperation. *International Economic Review* 32 (3): 551–559.

Grossman, S. J. 1981. Nash equilibrium and the industrial organization of markets with large fixed costs. *Econometrica* 49 (5): 1149–1172.

Holmstrom, B. 1982. Moral hazard in teams. *The Bell Journal of Economics* 13 (2): 324–340.

Hurter, A. P., and P. Lederer. 1986. Competition of firms: Discriminatory pricing and location. *Econometrica* 54 (3): 623–640.

Itoh, H. 1991. Incentives to help in multi-agent situations. *Econometrica* 59 (3): 611–636.

Katz, M. 1991. Game-playing agents: Unobservable contracts as precommitments. *Rand Journal of Economics* 22 (3): 307–328.

Koçkesen, L., and E. A. Ok. 2004. Strategic delegation by unobservable incentive contracts. *Review of Economic Studies* 71: 397–424.

Kwon, S. Y. 1996. The impact of competition within the client's industry on the auditor selection decision. *Auditing: A Journal of Practice & Theory* 15 (1): 53–69.

Lennox, C., and C. W. Park. 2003. Accounting alumni, audit firm appointments, and audit committee characteristics. Working paper, Hong Kong University of Science & Technology.

Liu, X. H. 2002. Profit sharing in audit partnerships. Ph.D. dissertation, The University of British Columbia.

Narayanan, V. G. 1995. Moral hazard in repeated professional partnerships. *Contemporary Accounting Research* 11 (2): 895–917.

O'Keefe, T., D. A. Simunic, and M. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Pae, S., and S. Yoo. 2001. Strategic interaction in auditing: An analysis of auditor's legal liability, internal control system quality, and audit effort. *The Accounting Review* 76 (3): 333–356.

Reitman, D. 1993. Stock options and the strategic use of managerial incentives. *American Economic Review* 83 (3): 513–524.

Schwartz, R. 1997. Legal regimes, audit quality, and investment. *The Accounting Review* 72 (3): 385–406.

Sklivas, S. 1987. The strategic choice of managerial incentives. *Rand Journal of Economics* 18 (3): 452–458.

Trompeter, G. 1988. An investigation into the effect of competition, partner compensation, legal liability, and authoritative pronouncements on auditor judgment and independence. Ph.D. dissertation, The University of Wisconsin–Madison.

———. 1994. The effect of partner compensation schemes and generally accepted accounting principles on audit partner judgment. *Auditing: A Journal of Practice & Theory* 13 (2): 56–68.

U.S. Senate, Committee on Government Operations, Subcommittee on Reports, Accounting, and Management. 1977. *The Accounting Establishment: A Staff Study.* 95th Cong., 1st sess. S. Doc. 95–34. Washington, D.C.: Government Printing Office.

Who Audits America. 2003. 50th edition. Menlo Park, CA: Data Financial Press.

Zeff, S. 2003. How the U.S. accounting profession got where it is today: Part II. *Accounting Horizons* 17 (4): 267–286.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 703–722

Knowledge, Adaptivity, and Performance in Tax Research

Anne M. Magro
University of Oklahoma

ABSTRACT: The ability to adapt decision making to the features of decision tasks and contexts is likely an important component of professional decision making, but not all professionals will exhibit this adaptivity. One explanation for decision makers' failure to adapt when adaptivity is appropriate is that they do not possess relevant knowledge of the features of decision tasks and contexts, or *institutional knowledge*. In this study, I examine the relations between institutional knowledge, information search adaptivity, and performance using an experimental research design in the tax decision-making setting. The results of the study are consistent with predictions. Tax professionals with relevant institutional knowledge responded to the differential features of the tax-planning and compliance contexts by conducting broader and more extensive information search in planning than in compliance; professionals lacking institutional knowledge did not exhibit such information search adaptivity. In addition, tax research performance increased with information search adaptivity, and that adaptivity mediated the relation between institutional knowledge and tax research performance. This study extends the accounting and psychology adaptivity literatures and contributes to tax practice and education.

Keywords: *information search adaptivity; performance; knowledge; tax decision making.*

Data Availability: *Contact the author.*

I. INTRODUCTION

Features of accounting decision tasks and contexts vary, and tailoring the decision process to those features is an appropriate adaptive strategy. However, this *adaptivity* can occur only if decision makers know that decision tasks and contexts differ in relevant ways *and* possess appropriate strategies to deal with those differences (Payne et al. 993). Institutional knowledge, or knowledge of the features of tasks and contexts in the

---

This paper is based on my dissertation completed at the University of Illinois at Urbana–Champaign. I thank my dissertation committee (Eugene Willis, Ira Solomon, Dan Stone, and Janet Sniezek) for their support and assistance throughout the project. I also appreciate the helpful comments of participants in research forums at the University of Illinois at Urbana–Champaign, University of Missouri, Case Western Reserve University, Rutgers University, and University of Waterloo, and participants at the 1999 ABO Research Conference. The comments of B. Cloyd, H. Finstuen, J. Hammersley, K. Kadous, J. Kennedy, L. Koonce, M. Lipe, M. Roberts, M. Shelley, and two anonymous referees are especially appreciated. This dissertation would not have been possible without the financial assistance of the University of Illinois Department of Accountancy and an Ernst & Young Dissertation Fellowship, as well as the access to subjects and computer resources provided by the participating firm.

Editor's note: This paper was accepted by E. Michael Bamber, Guest Editor.

Submitted January 2003
Accepted August 2004

professional decision-making setting, is critical to exhibiting the adaptivity that is necessary to perform well across a variety of accounting tasks and contexts. Therefore, I predict that decision makers who understand the differential features of decision tasks and contexts and whose decision strategies are based on an understanding of those features should perform better than do those without such knowledge. In this study, I examine (1) the relations between the features of an accounting decision setting and appropriate information search strategies, (2) the moderating role of institutional knowledge on information search adaptivity, and (3) the relations of institutional knowledge, information search adaptivity, and performance in a professional accounting task: tax research.

Tax research includes the processes of issue identification, information search, authority selection and documentation, and conclusion formation. The results of the tax research process serve as the primary inputs to the advice preparation process. Tax professionals serve as client advisers and advocates in an environment characterized by significant exposure to practice risk (Kadous and Magro 2001), but they can reduce this exposure by relying on high-quality tax research. Because tax research is critical to tax decision making, it is important to understand how individual characteristics interact with features of tax decision contexts to drive tax research performance. Magro (1999) conducted a detailed task and context analysis and identified key differences between two primary tax decision-making contexts, tax planning and tax compliance.[1] She demonstrated that the planning context is characterized by greater ambiguity, complexity, and justifiability demands than is the compliance context. While tax professionals conduct information search using the same tools and information sets in both contexts, tax professionals in the planning context should respond to these contextual differences by conducting broader and more extensive information search than do professionals in the compliance context. I predict, however, that institutional knowledge acquired through experience with the contexts moderates the relation between context and adaptivity. Specifically, I predict that tax professionals who do not possess relevant institutional knowledge will not exhibit the predicted information search adaptivity. Further, I predict that the ability to appropriately adapt information search to the context is a characteristic of good performance in tax research and that adaptivity mediates the relation between knowledge and performance in this setting.

To test these hypotheses, I conducted an experiment using tax staff and seniors from an international accounting firm as participants. Decision context was manipulated within participant, and institutional knowledge was measured. As predicted, those with institutional knowledge exhibit the predicted adaptivity by conducting broader and more extensive information search in planning than in compliance, but those without such knowledge did not. Tax research performance increased with adaptivity, and adaptivity mediated the relation between institutional knowledge and performance. That is, those who knew that planning is characterized by greater ambiguity, complexity, and justifiability demands than is compliance conducted broader and more extensive information search in planning, and that information search adaptivity contributed to better performance on the overall tax research task.

This study makes three main contributions to the literature. First, the study demonstrates that task, context, and individual characteristics combine to drive information search strategies in a professional setting and that a detailed analysis of the setting is necessary to predict when and how decision makers will adapt their decision strategies. Second, this study is the first to focus on the importance of institutional knowledge to decision-making process and performance in accounting. Finally, the results of this study contribute to tax practice. The results suggest that information search adaptivity is an important component of tax research performance and mediates the relation between institutional knowledge and performance. The results of this study suggest that educators and accounting firms can improve the education and training of tax professionals by providing opportunities for them to acquire institutional knowledge, including knowledge of both task and context features and appropriate adaptive strategies.

---

[1] Tax planning occurs when the client approaches the tax professional for advice before all facts are established and the opportunity exists to affect the way a transaction will be structured. Tax compliance occurs when a client approaches the tax professional after all facts have been established or when tax considerations will not affect the structure of a transaction.

In the next section, I develop my theory and hypotheses, and in Section III I describe my experimental method. Section IV contains the results and discussion, and the final section contains conclusions, limitations, and directions for further research.

II. THEORY

Contingent Decision Behavior

Adaptive Decision Making

Contingent decision behavior is a term used to describe the phenomenon by which decision makers employ varying decision strategies dependent on the situation (Payne et al. 1993; Petty and Cacioppo 1986; Beach and Mitchell 1978; Reder 1987; Siegler 1988). With their adaptive-decision-maker framework, Payne et al. (1993) make a significant contribution to our understanding of contingent decision making by drawing on a sizable body of research (see Payne et al. [1993] for a review) to develop a coherent theory that predicts and explains such behavior. Three principles of the adaptive-decision-making framework are that (1) decision strategies have different characteristics, (2) features of the decision problem (i.e., task and context features),[2] the social context of the decision, and the decision maker influence the anticipated costs and benefits of decision strategies, and (3) decision makers attempt to balance cost and benefit concerns when deciding how to make a decision. Thus, different individuals may employ different strategies for the same problem, or the same individual may select different strategies at different times. While this basic framework is valid across decision settings, the features of decision settings and measures of adaptivity considered in prior studies may not be meaningful in more naturalistic decision settings characterized by complex, unstructured tasks (Klein et al. 1993).

Features of a Professional Setting: Tax Decision Making

Accounting decision settings are characterized by features that distinguish them from the settings previously studied in the adaptivity literature, including well-educated professional decision makers, a multiperson decision-making environment, and the existence of a vast body of professional and legal standards that inform the decision process. The tax decision-making setting is further distinguished by the prominence of formalized information search. While important in many decision settings, information search is the primary focus of tax research, a fundamental process in tax decision making. The product of the tax research process (usually a memo with supporting documentation) serves as the principal input to the client advice-formation process, and forming advice for the client is generally the purpose of tax engagements.

Another important feature of the tax setting is that most decision making occurs in one of two decision contexts, tax planning or tax compliance. In planning, the client has not yet entered into a transaction and the opportunity exists to affect the way in which the transaction is structured. In compliance, the client has either completed the transaction in question or is unable or unwilling to alter how a planned transaction will be implemented. In her descriptive study of features of tax decision contexts, Magro (1999) shows that different information characteristics, goals, responsibilities, and potential costs between planning and compliance contexts cause planning problems, in general, to be characterized by greater ambiguity, complexity, and justifiability demands than are compliance problems. I extend the prior literature by developing and testing predictions about how differences across tax contexts and professionals manifest in adaptivity along specific information search dimensions and differential tax research performance.

---

[2] Task features that have been considered include task complexity, response mode, information display, and agenda effects. Context features include similarity, quality of the option set, reference points, and framing.

Ambiguity and complexity. *Ambiguity* is a state of uncertainty with respect to information, decision alternatives, constraints, and the criteria by which decisions will be judged (Beach and Mitchell 1978; Payne et al. 1993). In contrast to the compliance context, a specific transaction structure has not been implemented by the client in the planning context (in fact, the objective in tax planning is to determine what the facts should be in order to achieve the client's tax and economic goals), and the law as well as the client's characteristics, resources, and goals can change during the planning process (Magro 1999).[3] Such changes can expand the set of potential transaction structures, render previously appropriate structures unusable, or impose different criteria by which the decision will be evaluated. *Task complexity* increases with the number of problem components, the number of criteria on which the decision will be evaluated, and the decision's potential influence on future events (Beach and Mitchell 1978; Payne et al. 1993). Planning is characterized by more problem components and criteria by which decisions will be judged than is compliance. The tax professional's goals in planning are to identify transaction structures and tax treatments that satisfy both the client's overall economic and tax goals subject to statutory and client constraints. In compliance, the tax professional's goal is only to identify appropriate tax treatments taking into consideration the client's tax goals and statutory constraints (Sommerfeld and Streuling 1981; Causey and McNair 1990). Planning judgments, which determine the structure of transactions, also have greater potential to influence future events. For example, the choice of business entity may have significant long-term tax and economic effects for the client. Together these features make the planning context more ambiguous and complex than is compliance.

Justifiability demands. The need to support one's judgments and decisions is an important feature of the social context in which professional decision making occurs. Differences in practice risk (Kadous and Magro 2001) drive differences in the justifiability demands of tax planning and compliance (Magro 1999). Planning work exposes the tax professional to greater risk than does compliance work because compliance advice is subsumed by planning advice. The Statement on Standards for Tax Services (SSTS) No. 8 (AICPA 2000) states: "In advising or planning with a client on tax matters, the CPA should assume that the advice will affect the manner in which the matters or transactions considered ultimately will be reported on the client's tax returns. Thus, for all tax advice the CPA gives to a client, the CPA should follow the standards in SSTS No. 1 relating to tax return positions." As such, a professional who gives bad advice in tax planning may be subject to all of the costs associated with compliance work[4] plus the additional civil liability and loss of reputation associated with planning malpractice litigation. This additional civil liability results because planning advice is more significant for the client than is compliance advice—clients may incur significant costs to implement tax planning advice and such implementation is likely irreversible with long-term effects that go far beyond the scope of taxes (Quattrochi 1982). Compliance advice neither imposes the same potential costs nor has the same potential impact on the client.

---

[3] For example, client facts would change if an S-corporation requested advice, but before implementing a recommended transaction structure, lost its S-corporation status. Goals would change if a client intended to finance an acquisition with cash but other needs for that cash arose, leading them to consider alternative means to finance or alternative structures like exchanges or installment sales to effect the acquisition.

A consideration of educational materials, professional assessments of contextual features, and tax malpractice insurance activity supports the contention that tax professionals are more vulnerable in planning than in compliance. Tax research textbooks identify planning work as riskier for the professional (Karlin 2003) and recommend using disclaimers when communicating tax-planning advice to clients (Sommerfeld and Streuling 1981). Experienced tax managers in an experimental setting assess the justifiability demands of planning problems as higher than those of compliance problems (Magro 1999). While more malpractice insurance claims are filed in compliance than in planning,[5] tax professionals are exposed to greater expected costs in planning because the costs of settling claims is greater.[6] In addition, more malpractice lawsuits are filed in planning, leading to higher litigation costs (Ferguson 1996), and cases that make it to trial are likely those that pose the greatest risk to the tax professional's reputation. In the next subsection, I address the moderating role of institutional knowledge in the relation between these contextual differences and appropriate adaptive responses.

Institutional Knowledge

Payne et al. (1993) argue that individuals will not exhibit adaptive-decision-making behavior if they fail to (1) identify differences between the decision problems or contexts, (2) recognize that an adaptive response to identified differences is appropriate, or (3) generate appropriate adaptive strategies in response to identified differences. Thus, adaptivity is only expected when the decision problems or contexts differ and the decision maker possesses knowledge of relevant differences, the appropriateness of an adaptive response, and appropriate adaptive strategies. As such, any test of adaptivity is really a joint test of the three conditions.

The importance of knowledge to adaptive decision making is not surprising as knowledge is well established as a key component of performance (Libby and Luft 1993; Libby 1995). Types of knowledge relevant to accounting decision making include technical knowledge (Cloyd 1995b, 1997; Spilker 1995), tacit knowledge (Libby and Tan 1997), client and general business knowledge (Bell et al. 1997), and institutional knowledge (Magro 1999). Particularly important in the tax setting are knowledge of the business setting, knowledge of the client, technical tax knowledge (i.e., knowledge of the tax law as it applies to specific transactions), and institutional tax knowledge (i.e., knowledge of the tax decision-making environment). While each of these types of knowledge likely has different effects on tax professionals' judgments and decision making, prior research in tax has focused largely on technical tax knowledge. Possession of such knowledge has been shown to have a positive effect on identification of tax issues (Bonner et al. 1992), assessment of distinguishing features of tax cases (Davis and Mason 2003), recommendations of reporting positions (Cloyd 1995a; Roberts and Klersey 1996), information search under accountability demands (Cloyd 1995b, 1997), and keyword selection (Spilker 1995).

---

[4] In the compliance context, the tax professional may (1) be subject to preparer penalties (generally in the $250–$1,000 range) if she completes the return, (2) be liable for client interest or client penalties (usually a percentage of the additional tax due), and (3) face loss of reputation.

[5] Although almost half of the tax claims to the AICPA-endorsed professional liability insurer are for compliance issues such as math or filing errors (Stevens 2000), these constitute frequent, but relatively low-cost, claims that are usually resolved out of court (Kahan 1999). Such claims likely are not associated with great loss of reputation, which tax professionals cite as the most important cost they face (Collins et al. 1990).

[6] In a personal communication, Joseph Wolfe, Director of Risk Management for the AICPA-endorsed professional insurance program, indicates that the average cost of defending and settling claims is about three times greater for tax-planning than for tax compliance claims. Tax-planning claims are frequently settled for over $100,000, while most tax compliance claims cost less than $50,000 to defend and settle.

Several prior studies have investigated the effects of contextual features on tax decision making (cf., Cloyd 1995b, 1997; Kadous and Magro 2001; Alexander 2003), but only one prior study has investigated the effects of institutional knowledge on tax professionals' judgment and decision making. Spilker (1995) addressed the effects of declarative knowledge (operationalized as technical knowledge), procedural knowledge (operationalized as tax research skill), and time pressure on keyword selection. The possession of a domain-specific skill, such as the ability to use a tax database to complete tax research, is an element of knowledge about the decision setting and, as such, a type of institutional knowledge. Participants with both technical and institutional knowledge identified more relevant keywords than did those with only technical knowledge, regardless of time pressure, and their performance advantage increased in the presence of time pressure. Spilker's (1995) study demonstrates that institutional knowledge can be an important component of performance in a tax task and that technical and institutional knowledge can have distinct effects on performance.

An important element of institutional knowledge in tax is knowledge of the differential contextual features of planning and compliance. Tax research begins with the identification of issues based on the client's facts. Using these issues as a starting point, the tax professional conducts information search to identify authorities that inform these issues and to meet professional and statutory requirements. This process is iterative with the set of issues and relevant facts refining or expanding during the information search process. From the authorities accessed during information search, the tax professional selects authorities to include in the communication of conclusions.

Because tax planning is characterized by greater ambiguity, complexity, and justifiability demands than is compliance, the tax professional has incentives to employ different information search strategies in planning and compliance to achieve desired levels of decision quality. Two relevant dimensions of information search in tax are *extent* and *breadth*,[7] where extent describes ***how much*** search is done and breadth describes ***which authorities*** are accessed during search. The authorities accessed during search can be described along several dimensions including the content of the authority, the level of authoritative support provided by the authority, or the type of authority. Specific contextual differences across planning and compliance are more likely to result in differences in one dimension of breadth than another. Specific predictions with respect to the relations between ambiguity, complexity, and justifiability demands and information search breadth and extent are developed below.

Planning is more ambiguous and complex than is compliance and is therefore characterized by more problem components and decision criteria and by greater uncertainty and potential impact on future events. In response to these differences, tax professionals should access authorities with different content and level of authority during information search. First, the professional should access a set of authorities that address multiple issues and transaction structures.[8] Further, tax professionals can use secondary sources[9] that summarize various types of primary authorities to gain a better overall understanding of the law without accessing those sources directly. In the compliance context, these types of search have less value because the decision problem is delineated with respect to one transaction and its related issues, and the tax professional knows that the facts and relevant law will not change.

---

[7] Some prior research on information search in accounting characterized information search as either directed or sequential (Biggs et al. 1988). While information search in tax could be described similarly, there is no theoretical basis for predicting differences in the directedness of search across the planning and compliance contexts.

This analysis suggests that ambiguity and complexity differences affect the *focus* of search across planning and compliance contexts by making a broad search with respect to the content of the authorities and their level of authority more valuable in planning. Decision makers, however, cannot exhibit such information search adaptivity without knowledge of the contextual differences that make the adaptivity appropriate. The moderating role of institutional knowledge in the relation between contextual features and information search adaptivity is formalized in the following hypothesis.

> **H1:** Tax professionals who know (do not know) that the planning context is characterized by greater *ambiguity and complexity* than is the compliance context will (will not) conduct *broader* information search in planning than in compliance.

A similar argument can be made with respect to justifiability demands in planning and compliance. Given the increased exposure to practice risk in the planning context, the tax professional should make greater efforts to identify plans with high levels of support and to accurately assess the risks of those plans to satisfy the increased justifiability demands. To meet statutory and professional support requirements, the tax professional is required to consider all types of primary authorities (i.e., IRC sections, Regulations, cases, Revenue Rulings, etc.) and to identify all relevant authorities before advising the client. The consequences of failing to do so are greater in planning than in compliance, so the tax professional has an incentive to conduct both a more extensive search by considering *more* authorities (to meet requirements to identify all relevant authorities) and a broader search by considering more *types* of authorities in planning (to meet requirements to consider all primary authority types). While the increased ambiguity and complexity of planning is predicted to result in the use of more secondary sources that summarize multiple primary authorities, the increased justifiability demands of the planning setting should lead the tax professional to directly access the primary sources as well. Tax professionals, however, cannot respond to the increased justifiability demands in planning unless they know that such demands exist. The prediction that institutional knowledge moderates the relation between the context and information search adaptivity leads to the following hypothesis.

> **H2:** Tax professionals who know (do not know) that the planning context is characterized by greater *justifiability demands* than is the compliance context will (will not) conduct *broader and more extensive* information search in planning than in compliance.

---

[8] For example, if a taxpayer wished to transfer stock to her child, she could structure the transaction as a gift, a sale, an exchange, or a bequest. Each of these would be considered a different transaction structure.

[9] Secondary sources are those that provide no authoritative support in litigation and cannot be relied upon when assessing the strength of support for a position. Primary authorities are those upon which the tax professional may rely.

Tax Research Performance

The tax research process is important because the products of that process (often a memo and supporting documentation in the multiperson decision setting) serve as the primary inputs to the advice-formation process. Documentation of the research process is also critical when defending the firm in litigation or against challenges by the IRS. Tax managers and partners at an international accounting firm report that, when evaluating tax research quality, they consider the (1) quality of the underlying information search process (comprehensiveness, efficiency, and authorities selected), (2) quality of authority analysis, (3) quality of documentation, (4) strength of arguments, (5) level of authoritative support for positions recommended, and (6) quality of the writing (interviews at an international accounting firm).

Information search adaptivity in response to contextual differences should be associated with tax research performance because it impacts the quality of the underlying information search process and the level of authoritative support for the positions recommended. For example, failing to conduct an extensive search in the planning context would affect the perceived comprehensiveness of information search. Further, failing to conduct a broad search in the setting would impact the specific authorities selected (a characteristic of the quality of the search) and thus the authoritative support for the positions recommended. As a key determinant of information search adaptivity, institutional knowledge of contextual features should have an indirect effect on tax research performance through its effect on information search adaptivity.

H3: Tax research performance increases with information search adaptivity.

H4: The effect of institutional knowledge of ambiguity, complexity, and justifiability demands differences across planning and compliance contexts on tax research performance is mediated by information search adaptivity.

III. RESEARCH METHOD

Task

The experimental instrument contained two tax research assignments. Each assignment provided information about a different client and included a brief history of the client's business, a discussion of pertinent relationships, and a description of a completed or planned transaction. Participants were instructed to conduct tax research, document those sources they wished to submit, and convey the results of their research in a memo to the tax manager. Participants were allowed to use only the CCH ACCESS database for research and were limited to two hours to complete each assignment (total experiment length was 4.5 hours including the post-experiment questionnaire).[10] Participants completed the research cases while working at their own computers during a designated block of time. Two administrations of the experiment were conducted. As analysis of the data shows no effect of group, they are combined in all further analyses.

Participants

Twenty-eight staff- and senior-level tax professionals employed by an international accounting firm participated in the experiment. Tax professionals at this level complete a significant majority of the tax research assignments in the firm. Participants were volunteers recruited by a firm partner. Participants' tax experience ranged from eight to 42 months with a mean (SD) of 16.9 (6.9) months. Nineteen participants had advanced degrees, including 12 J.D.s. Demographic data suggest that participants spend a significant proportion of their chargeable hours on tax research, are familiar with the CCH ACCESS database, and conduct research related to both planning and compliance problems.[11] Table 1 presents selected demographic data by high and low institutional knowledge group.

---

[10] Based on interviews at an international accounting firm, initial investigations of research issues generally last three to four hours. I limited research to two hours for each project due to participant availability constraints.

Incentives

Incentives were provided both to encourage participation and to motivate participants to take the project seriously. To encourage participation in the study, volunteers received training hours for their participation and were entered into a lottery for two prizes (a pair of tickets to see a professional basketball game or $250). The lottery was run separately for each administration of the experiment. To motivate them to take the project seriously, participants were informed that firm managers would evaluate their research products (research memo and supporting documentation) and that 25 percent of the participants would be contacted later to explain their work. The possibility of being held accountable for their work products is representative of the natural setting in which staff submit their work products to a supervisor who evaluates them. To maintain the integrity of the experimental

TABLE 1
Mean (SD) Demographic Variables

Panel A: Knowledge of Ambiguity and Complexity Differences

	AMB/COMP_KNOW[a]		*JUST_KNOW*[b]	
	High n = 18	Low n = 10	**High** n = 10	Low n = 18
Experience (months)	17.64	14.40	20.40	14.28
	(8.100)	(3.096)	(9.876)	(3.012)
J.D. (percentage of participants)	50.00	30.00	60.00	33.33
Percentage of time on planning work	57.33	23.00***	58.20	37.78
	(23.522)	(20.575)	(23.218)	(27.930)
Percentage of time on compliance work	34.89	57.00**	33.80	47.78
	(19.041)	(15.312)	(18.558)	(20.380)
Exposure to partnership problems (percentage of participants)	55.56	40.0	50.00	50.00
Exposure to S-corporation problems (percentage of participants)	50.0	40.0	50.00	44.44
Exposure to both technical topics (percentage of participants)	44.44	40.0	40.00	44.44

*, **, *** $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

[a] *AMB/COMP_KNOW* is the measure of institutional knowledge of ambiguity and complexity differences across planning and compliance.

[b] *JUST_KNOW* is the measure of institutional knowledge of justifiability demand differences across planning and compliance.

---

[11] Participants reported that they spend, on average, 34.1 percent of their chargeable hours on tax research. Reports ranged from 5 percent to 90 percent, with the two most senior participants reporting the lowest percentages. Participants reported that they use the CCH ACCESS database for 86 percent of their tax research assignments on average and that their total tax research time is allocated between planning (53 percent), compliance (41 percent), and other research (6 percent).

instructions, managers actually evaluated staff work products, and I contacted seven participants to request clarification or additional information.

Independent Variables

Context

Context was manipulated within-participant at two levels (planning and compliance) by changing the timing of the client's request for advice. In the compliance condition, the transaction had already occurred while in the planning condition the client was considering the transaction. The planned transaction had negative tax effects that could be avoided by restructuring the transaction. Each participant completed one research assignment that involved the sale of a partnership interest and one that addressed a distribution by an S-corporation to a shareholder. Half of the participants did the partnership problem in the planning context and the S-corporation problem in the compliance context, and the other half did the partnership problem in the compliance context and the S-corporation problem in the planning context. The order of the problems was manipulated, yielding four conditions. Order is not significant, and the data are collapsed across order groups in all reported analyses. Whether in the planning or compliance contexts, the materials identified only one potential structure for the transaction.

Institutional Tax Knowledge

I created three dichotomous variables (*AMB_KNOW*, *COMP_KNOW*, and *JUST_KNOW*) representing the participant's knowledge of ambiguity, complexity, and justifiability demand differences across tax planning and compliance contexts. In the post-experimental questionnaire, participants were asked to assess the complexity, ambiguity, and justifiability demands of planning and compliance problems. Each feature was defined[12] then immediately followed by a question of the form: "In general, how ambiguous are compliance problems?"[13] For the ambiguity question, the endpoints of the seven-point Likert response scale were anchored by "Not at all Ambiguous" and "Completely Ambiguous." These questions were embedded in an 82-item questionnaire that contained 59 items on which they responded using Likert scales. The three questions related to planning were grouped together and were separated by approximately 20 questions from the group of three compliance questions.

To control for the fact that individuals may use Likert scales differently, participants also responded to a question at the end of the questionnaire that was designed to identify how they used the scale. Participants were asked to place a vertical line through the point on the scale that delineated the smallest meaningful difference from a designated point.[14] Institutional knowledge was considered high (value of 1) when the difference between the participant's ratings for the two contexts exceeded this self-reported meaningful difference on the rating scale. I combined *AMB_KNOW* and *COMP_KNOW* to form *AMB/COMP_KNOW*, which takes on a value of 1 when either *AMB_KNOW* or *COMP_KNOW* is high.[15] Using these methods, 18 participants were classified as high *AMB/COMP_KNOW* and ten were classified as high *JUST_KNOW*.[16] I test H1 and H2 with *AMB/COMP_KNOW* and *JUST_KNOW*.

---

[12] I did not define the terms "planning" and "compliance" for participants because they are introduced in all leading introductory tax texts, and tax professionals generally share a uniform definition of these terms. To the extent that the participants in this study did not share a common definition of these terms, noise but not bias is introduced into the data.

[13] I use institutional knowledge measures based on assessments of planning and compliance contexts in general rather than using questions assessing the specific planning and compliance cases used in the study because, while the perceived features of the specific problems being researched might affect information search behaviors, information search strategy in initial research is likely driven by *ex ante* beliefs about the features of the contexts in general. Further, the cases were carefully selected to be similar in ambiguity, complexity, and justifiability demands, so any assessed differences across case by the participants should be attributable to differences in context. Results are qualitatively the same using knowledge measures constructed from case-specific questions.

To test the mediating role of institutional knowledge in the relation between adaptivity and performance predicted in H4, I created a measure of overall institutional knowledge, *INST_KNOW*, by summing *AMB/COMP_KNOW* and *JUST_KNOW*. Because *AMB/COMP_KNOW* and *JUST_KNOW* each take on a value of 0 or 1, *INST_KNOW* takes on a value of 0, 1, or 2.

Information Search Adaptivity

I developed an overall measure of adaptivity to test H3 and H4, which address the effects of information search adaptivity and institutional knowledge on performance. I calculated z-scores for each of the five search dimensions used as the dependent variables in tests of H1 and H2 (described below) and computed the difference for each across the planning and compliance contexts. I then summed the difference scores across the five dimensions to yield the variable *ADAPTIVITY*.

Dependent Variables

Information Search Adaptivity

The CCH ACCESS PATH function keeps a complete record of all research steps on a project, and its printout cannot be altered without detection. Using the participants' PATHs, I compiled a comprehensive list of sources accessed by the participants for each problem. Participants, in total, identified 354 sources in one problem and 283 sources in the other.[17] I classified each source along the following dimensions: (1) transaction structure addressed, (2) type of authority (e.g., IRC section, Regulation, Revenue Ruling, or CCH Explanation), and (3) issue addressed. A second rater with ten years' tax experience classified each source with respect to the subjective measures, transaction structure, and issue, for a sample of 127 of the 637 authorities. Each of the sources was classified without rater knowledge of the condition in which it was generated or the subject who identified it. Inter-rater reliability averaged 88.9 percent and ranged from 87.3 percent to 91.2 percent.[18]

---

[14] Participants received a seven-point Likert scale anchored by 1 and 7 with a vertical line (labeled "A") drawn through the scale at 3. The instructions read, "You answered several questions using scales like the one below. To help us understand how you used this scale, please place a vertical line at the point that marks the smallest important difference in ratings from the indicated line 'A.' That is, when evaluating your responses to the questions on the preceding pages, we should consider the ratings of two previous items significantly different if the distance between your marked lines on those items is at least as great as the distance between line 'A' and your mark on the scale below."

[15] While raw ambiguity and complexity differences between planning and compliance varied within subject, all subjects with high *AMB_KNOW* also had high *COMP_KNOW*. *AMB_KNOW* and *COMP_KNOW* were combined because ambiguity and complexity are predicted to have similar effects on the search dimensions measured in this study.

[16] One participant in each low-knowledge group actually assessed greater ambiguity and complexity or justifiability demands for compliance than for planning. If these participants are reclassified as high-knowledge with their results reverse-scored, the results are qualitatively the same.

[17] I count as a source the individual units available through CCH ACCESS. For example, one paragraph of an Internal Revenue Code section, multiple paragraphs of an Internal Revenue Code section, a case, or an entire tax bill each would be designated as one source if CCH has created the access to the source in that size chunk.

Search breadth is a multidimensional theoretical construct that was operationalized in this study as number of issues, number of transaction structures, percentage of secondary sources, and number of authority types. While other measures of search breadth in tax may exist, these measures are particularly suited to the contextual features manipulated herein. *Number of issues* and *number of transaction structures* were measured as the number of unique issues and transaction structures, respectively, addressed in the sources accessed by a participant. *Percentage of secondary sources* was calculated by determining the authority type of each source, coding the sources as primary or secondary based on definitions in tax texts, and calculating the percentage of secondary source authorities in each participant's authority set. This percentage was transformed using the arcsin transformation for statistical testing. Percentage rather than number of secondary sources was used because it measures the emphasis on secondary sources rather than just the number of such sources. For example, one would interpret the fact that a participant accessed five secondary sources differently if the participant accessed a total of ten authorities than if he accessed 30. *Number of authority types* was measured as the number of different authority types accessed by a participant.

Extent of search is also a multidimensional theoretical construct but is operationalized here with a single measure, number of authorities accessed. While time of search would be another appropriate measure of the extent of search in tax, participants in this study were under significant time pressure and all used the full time available to them to complete the research. *Number of sources* was measured as the total number of sources accessed minus the number of repeat hits.

These search measures were used as the dependent variables to test the hypotheses relating context and institutional knowledge to adaptivity. Consistent with the theory, I test H1 with number of issues, number of transaction structures, and percentage of secondary sources. Number of authority types would not be an appropriate measure of breadth here because increased complexity would not increase the value of considering both Revenue Rulings and case law and increased ambiguity would be addressed by considering more secondary sources that summarize multiple authority types. Consistent with the theory with respect to justifiability demands, I test H2 with number of authority types (breadth of search) and number of authorities (extent of search).

Tax-Research Performance

The performance measure for H3 and H4 is a continuous variable (*QUALITY*) based on the *average* assessed quality of each participant's tax-research products (i.e., the planning and compliance memos). The tax manager (a CPA with a J.D., an M.S. in accounting, and an L.L.M. in tax) was from the same accounting firm as the participants and had significant accounting (14.6 years) and tax (6.1 years) experience. A second rater also evaluated the research products of the participants. This rater was a doctoral student with 17 years' tax experience who had been a Big 5 tax manager and a sole practitioner. Both were unaware of the research hypotheses, and evaluations were blind. The two evaluators' assessments are highly correlated ($r = 0.901$). Reported analyses are based on the tax manager's assessments.

---

[18] Inter-rater agreement on the S-corporation problem was 87.3 percent for issues and 87.3 percent for transaction structures. Inter-rater agreement for the partnership problem was 89.7 percent for issues and 91.2 percent for transaction structures.

Each evaluator received a packet containing the following: (1) instructions and demographic questions, (2) research assignments, completed memos, and supporting documentation, and (3) pre-numbered rating sheets. The instructions included a description of the conditions under which the staff members were to complete the research and instructions to "evaluate the staff member's research as if you had made the assignment and the staff member had submitted the work for your use." The memos were grouped by decision context and technical topic (e.g., all partnership compliance memos were grouped together) with memos arranged randomly within each group. The evaluators rated quality on rating sheets that included a 101-point response scale (anchored at 0 by "Lowest Possible Quality" and at 100 by "Highest Possible Quality") and space for additional comments or rationale.

Research Cases

To control for differences other than context that could cause information-search variance across research assignments, I selected two research assignments from those used in Magro (1999). Magro (1999) developed planning and compliance versions of six research assignments based on the *Price Waterhouse Case Studies in Tax* (Jones 1994). In her study, 20 managers assessed the ambiguity, complexity, and justifiability demands of six cases in either the planning or compliance context. The two cases used in the present study were chosen from this set based on managers' assessments that the two cases were not significantly different on ambiguity, complexity, and justifiability demands within context (lowest $p = 0.242$). While managers' assessments of these features did not differ, the participants in this study might perceive the features of the problems differently. As such, I also considered the participants' assessed ambiguity, complexity, and justifiability demands of the specific research problems in the study. Similar to the results for the managers, participants' assessments of the features did not differ within context for the partnership and S-corporation problems (smallest $p = 0.417$). Means by problem and context for each group are presented in Table 2. As the assessments of ambiguity, complexity, and justifiability demands do not vary with assignment topic, differences in information search between the planning and compliance contexts in this study can be attributed to differences in the complexity, ambiguity, and justifiability demands between the contexts themselves and not to differences in the characteristics of the specific assignments.

Manipulation Check

A manipulation check for the context variable (planning/compliance) was included in the post-experimental questionnaire. Participants were asked two questions of the form, "In the (partnership) assignment, did the sale of the partnership interest occur before the client sought your firm's advice?" An answer of "No" in the planning condition and "Yes" in the compliance condition was interpreted as a successful manipulation. All participants answered both questions correctly. In addition, participants' assessments of ambiguity, complexity, and justifiability demands for the planning cases exceeded those for the compliance cases (all $p < 0.001$).

IV. RESULTS AND DISCUSSION

Hypothesis 1

In H1, I predict that tax professionals who possess the institutional knowledge that planning is characterized by greater ambiguity and complexity than is compliance will conduct a broader search in planning than in compliance, and those without such knowledge will not. Results of a repeated measures MANOVA (Khattree and Naik 1999) with context and *AMB/COMP_KNOW*, the measure of institutional knowledge, as independent variables

TABLE 2
Mean (SD) Assessments of Context Features by Technical Topic

Panel A: Managers (Magro 1999)

	Context	Partnership	S-Corporation	t
Ambiguity	Planning	4.14 (1.047)	4.49 (1.288)	$t_9 = 0.91$
	Compliance	2.79 (0.990)	3.06 (0.735)	$t_9 = 1.25$
Complexity	Planning	5.28 (0.755)	5.29 (0.731)	$t_9 = 0.03$
	Compliance	4.22 (1.451)	4.30 (0.906)	$t_9 = 0.15$
Justifiability Demands	Planning	4.76 (0.985)	4.75 (0.997)	$t_9 = -0.04$
	Compliance	4.28 (1.293)	4.05 (0.826)	$t_9 = -0.57$

Panel B: Tax Staff and Seniors

	Context	Partnership	S-Corporation	t
Ambiguity	Planning	5.01 (0.901)	4.92 (0.690)	$t_{26} = -0.28$
	Compliance	3.94 (0.813)	3.71 (1.333)	$t_{26} = 0.56$
Complexity	Planning	5.03 (0.863)	5.14 (0.676)	$t_{26} = 0.39$
	Compliance	4.23 (0.884)	4.13 (1.206)	$t_{26} = 0.25$
Justifiability Demands	Planning	5.53 (0.997)	5.69 (0.465)	$t_{18.4} = 0.56$
	Compliance	4.91 (0.753)	4.63 (1.016)	$t_{26} = 0.82$

*, **, *** $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

and number of issues, number of transaction structures, and percentage of secondary sources as dependent variables suggest that knowledge and context have an interactive effect on breadth of search ($F_{1,26} = 4.42$, $p = 0.045$). Tests of simple effects for the breadth measures show that in all cases search is broader in planning than in compliance for the high-knowledge group but not for the low-knowledge group. Descriptive statistics are presented in Table 3. For the high-knowledge group, search is broader in planning than in compliance for number of issues ($t_{17} = 4.48$, $p < 0.001$), number of transaction structures ($t_{17} = 3.79$, $p = 0.002$), and percentage of secondary sources ($t_{17} = 3.72$, $p = 0.002$). For the low-knowledge group, search is not broader in planning than in compliance for number of issues ($t_9 = 0.62$, $p = 0.550$), number of transaction structures ($t_9 = 0.80$, $p = 0.445$), or percentage of secondary sources ($t_9 = 0.61$, $p = 0.558$).

Hypothesis 2

In H2, I predict that search will be broader and more extensive in planning than in compliance when people know that planning is characterized by greater justifiability demands than is compliance. Without such knowledge, planning and compliance search will

TABLE 3
Breadth of Search by Knowledge of Ambiguity and Complexity (*AMB/COMP_KNOW*[a])

Panel A: Number of Issues

	Planning	**Compliance**	**t**
High *AMB/COMP_KNOW* (n = 18)			
Mean	13.11	6.83	$t_{17} = 4.48$***
(SD)	(6.07)	(3.95)	
Low *AMB/COMP_KNOW* (n = 10)			
Mean	14.80	13.70	$t_9 = 0.62$
(SD)	(6.03)	(4.30)	

Panel B: Number of Transaction Structures

	Planning	**Compliance**	**t**
High *AMB/COMP_KNOW* (n = 18)			
Mean	7.78	3.83	$t_{17} = 3.79$**
(SD)	(4.82)	(2.07)	
Low AMB/COMP_KNOW (n = 10)			
Mean	8.60	7.30	$t_9 = 0.80$
(SD)	(4.74)	(2.79)	

Panel C: Percentage of Secondary Sources

	Planning	**Compliance**	**t**
High *AMB/COMP_KNOW* (n = 18)			
Mean	.54	.24	$t_{17} = 3.72$**
(SD)	(.13)	(.11)	
Low *AMB/COMP_KNOW* (n = 10)			
Mean	0.61	.56	$t_9 = 0.61$
(SD)	(.04)	(.09)	

---

*, **, *** $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

[a] *AMB/COMP_KNOW* is the measure of institutional knowledge of ambiguity and complexity differences across planning and compliance.

not differ. Here, breadth is measured by the number of authority types, and extent is measured by the number of sources accessed. Descriptive statistics by *JUST_KNOW* are presented in Table 4. A repeated measures ANOVA with *JUST_KNOW* and context as independent variables reveals a significant knowledge by context interaction for breadth of search ($F_{1,26} = 6.55$, $p = 0.017$). The high-knowledge group used significantly more authority types in planning than in compliance ($t_9 = 3.34$, $p = 0.009$), and the low-knowledge group did not ($t_{17} = 0.81$, $p = 0.428$). Similar results obtain for extent of search. The repeated measures ANOVA reveals a significant knowledge by context interaction for extent of search ($F_{1,26} = 6.01$, $p = 0.021$). The high-knowledge group accessed significantly more authorities in planning than in compliance ($t_9 = 4.37$, $p = 0.002$), but the low-knowledge group did not ($t_{17} = 1.55$, $p = 0.139$).

Hypotheses 3 and 4

In H3, I predict that information search adaptivity (in the form of broader and more extensive search in planning than in compliance) is a characteristic of performance in the tax research task. To test this prediction, I regressed the summary measure of adaptivity, *ADAPTIVITY*, on the performance measure, *QUALITY*. *ADAPTIVITY* is significant in the regression ($F_{1,26} = 10.60$, $p = 0.003$, $R^2 = 0.29$)—performance in the research task increased with information search adaptivity.

TABLE 4
Breadth and Extent of Search by Knowledge of Justifiability Demands (*JUST_KNOW*[a])

Panel A: Number of Source Types

	Planning	Compliance	t
High *JUST_KNOW* (n = 10)			
Mean	5.60	3.00	$t_9 = 3.34^{**}$
(SD)	(2.46)	(1.33)	
Low *JUST_KNOW* (n = 18)			
Mean	4.61	4.22	$t_{17} = 0.81$
(SD)	(2.20)	(1.59)	

Panel B: Number of Sources

	Planning	Compliance	t
High *JUST_KNOW* (n = 10)			
Mean	25.40	13.80	$t_9 = 4.37^{**}$
(SD)	(9.66)	(7.10)	
Low *JUST_KNOW* (n = 18)			
Mean	22.83	19.61	$t_{17} = 1.55$
(SD)	(12.70)	(9.93)	

*, **, *** $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

[a] *JUST_KNOW* is the measure of institutional knowledge of justifiability demand differences across planning and compliance.

In H4, I predict that information search adaptivity mediates the relation between institutional knowledge and tax research performance. Baron and Kenny (1986) outline four conditions that must exist for mediation to be present. First, the initial variable (here, institutional knowledge) must be correlated with the outcome variable (here, performance). Second, the initial variable must be correlated with the mediator variable (here, adaptivity). Third, the mediator variable must be correlated with the outcome variable. Finally, when controlling for the effect of the mediator variable, the effect of the initial variable on the outcome variable either becomes insignificant (complete mediation) or decreases significantly (partial mediation). To test this predicted relationship, I ran regressions to test each of the four necessary conditions (Baron and Kenny 1986). The regressions are reported in Table 5.

A regression with tax research performance, as measured by *QUALITY*, as the dependent variable shows a significant effect of institutional knowledge, as measured by *INST_KNOW* ($t = 2.18$; $p = 0.039$). A regression with information search adaptivity, as measured by *ADAPTIVITY*, as the dependent variable shows a significant effect of *INST_KNOW* ($t = 2.71$, $p = 0.012$). A regression with *QUALITY* as the dependent variable shows a significant effect of *ADAPTIVITY* ($t = 3.26$, $p = 0.003$). In a regression with *QUALITY* as the dependent variable and *ADAPTIVITY* and *INST_KNOW* as the independent variables, *INST_KNOW* is not significant ($t = 0.96$, $p = 0.347$). Information search adaptivity completely mediates the relation between institutional knowledge and tax research performance (Baron and Kenney 1986).

Alternative Explanations

If technical knowledge, such as the knowledge of the rules governing partnerships, has the same impact on information search as does institutional knowledge and if technical

TABLE 5
Regressions for Mediation Analysis

Regression	Dependent Variable	Independent Variable(s)	Parameter Estimate	Standard Error	t-value	Pr > t
1	*QUALITY*[a]	*INST_KNOW*[b]	8.025	3.686	2.18	0.0387
2	*ADAPTIVITY*[c]	*INST_KNOW*	0.545	0.201	2.71	0.012
3	*QUALITY*	*ADAPTIVITY*	9.467	2.907	3.26	0.003
4	*QUALITY*	*INST_KNOW*	3.673	3.831	0.96	0.347
		ADAPTIVITY	7.983	3.297	2.42	0.023

[a] *QUALITY* is the average quality of the planning and compliance memo.
[b] *INST_KNOW* is the sum of the dichotomous variables *AMB/COMP_KNOW* and *JUST_KNOW*.
[c] *ADAPTIVITY* is the standardized sum of planning and compliance differences across the five measures of information search breadth and extent.

knowledge covaries with institutional knowledge, then failing to control for technical knowledge would make it impossible to disentangle the effects of the two types of knowledge.[19] While I did not directly measure technical knowledge of partnership and S-corporation law, I collected data with respect to technical-topic education and experience. Exposure to these technical topics through education or experience represents the opportunity to acquire technical knowledge (Bonner et al. 1992). Half the participants reported prior exposure to at least one of the topics in class, training, or practice.

I first consider the relation of exposure to the technical topics and the general institutional knowledge measures, *AMB/COMP_KNOW* and *JUST_KNOW*. The 24 frequency tables generated by crossing the institutional knowledge and exposure measures show no significant patterns (Fisher's Exact Test, lowest $p = 0.284$). The results of crossing institutional knowledge and exposure to the technical topics suggest that technical knowledge is not a correlated omitted variable in this study. To further assuage concerns that technical knowledge rather than institutional knowledge is driving the results, I ran repeated measures MANOVAs with exposure to the technical topic and context as the independent variables. For the MANOVAs using exposure to *either* technical topic, exposure does not interact with context for breadth of search ($F_{1,26} = 0.16$, $p = 0.690$) or extent of search ($F_{1,26} = -0.03$, $p = 0.869$). Similar results obtain for the MANOVAs using exposure to *both* technical topics as an independent variable: exposure does not interact with context for breadth of search ($F_{1,26} = 0.03$, $p = 0.863$) or extent of search ($F_{1,26} = 0.05$, $p = 0.831$).

V. CONCLUSION

This study explores the effects of task, context, and decision-maker characteristics on information search adaptivity and the relations between institutional knowledge, information search adaptivity, and tax research performance. The results are consistent with predictions and suggest that tax decision makers respond to features of the tax decision context when conducting information search. Consistent with Payne et al.'s (1993) theory that individual differences interact with task and context differences to drive adaptivity, those participants who possessed knowledge of differences in ambiguity, complexity, and justifiability demands across the planning and compliance contexts exhibited the predicted information search adaptivity while those without such knowledge did not. Those who knew that planning is characterized by greater ambiguity and complexity than is compliance investigated more transaction structures, more issues, and a greater percentage of secondary sources in planning than in compliance. Those who knew that planning is characterized by greater justifiability demands than is compliance used more authority types and accessed more authorities in planning than in compliance. Further, consistent with Libby and Luft (1993), institutional knowledge is associated with tax research performance as measured by the quality of tax research memos and supporting documentation. Finally, the ability to adapt information search to the context completely mediates the relation between institutional knowledge and tax research performance.

---

[19] Earlier, I addressed the issue of the relation between technical and institutional knowledge from a theoretical standpoint. Here, I report results with respect to the empirical relation between the two constructs in this participant group.

While I made theory-based predictions that the search of tax professionals without institutional knowledge would not differ across the contexts, I did not have theoretical grounds to predict what *level* of search breadth and extent they would exhibit. Interestingly, low-knowledge participants' search in both contexts looked most like high-knowledge participants' planning search. That is, low-knowledge participants conducted relatively broad and extensive searches in both contexts. The participants' assessed ambiguity, complexity, and justifiability demands are generally consistent with this information search pattern. If tax professionals conduct information search that is appropriate to planning problems while working in the compliance context, then both efficiency and effectiveness may be compromised. To the extent that search is more extensive than is appropriate for the compliance context, efficiency is impacted. This is particularly important as managers in Magro (1999) indicated that they would budget significantly less time for compliance research than for planning. Effectiveness may be compromised as well because broader search is different in focus, and such a search may focus on issues and transaction structures that are relevant in a planning context and fail to research relevant compliance issues and transaction structures in appropriate depth.

Limitations of this study provide opportunities for future research. While my participants had significantly more time to conduct tax research than did those in prior studies (two hours per problem), they had limited time relative to the natural context. Future research could investigate the effects of time pressure on information search across planning and compliance contexts. The measure of performance in this study was average overall quality of the tax research memos. Future research could investigate the various components of tax research quality and the role of information search therein. Ambiguity, complexity, and justifiability demands were manipulated together using a natural manipulation in the tax setting—planning versus compliance context. Within context, however, the characteristics may vary with details of the client's situation or the technical tax law. Future research could manipulate these characteristics within the planning or compliance context to determine if the relations between task and context features and adaptivity hold. I rely upon an indirect measure of technical knowledge in the supplemental analyses. Technical knowledge affects information search strategies (Cloyd 1995a, 1997; Spilker 1995), and the interaction of technical and institutional knowledge may yield information search strategies and forms of adaptivity not considered in the current study. Future research should consider the relative impacts of different knowledge types on information search strategies and tax research performance. Finally, planning and compliance contexts are also present in the financial accounting setting, but the relative risks and relevant features may differ therein. Future research could investigate the differences between tax and financial accounting planning and compliance contexts and the effect those differences would have on adaptivity.

In this study, I make three main contributions to the literature. First, I show that task, context, and individual characteristics combine to drive information search strategies in a professional setting and that a detailed analysis of the setting is necessary to predict when and how decision makers will adapt their decision strategies. Second, this study is the first to focus on the importance of institutional knowledge to decision-making process and performance in accounting. Finally, the results of this study contribute to tax practice. The results suggest that information search adaptivity is an important component of tax research performance (such adaptivity explained almost 30 percent of the variation in performance) and that adaptivity mediates the relation between institutional knowledge and performance. Universities and accounting firms can use this information to improve the education and training of tax professionals by providing opportunities for them to acquire institutional knowledge, including knowledge of both task and context features and appropriate adaptive strategies.

REFERENCES

Alexander, R. 2003. The effect of source credibility on tax professional judgment in consulting engagements. *The Journal of the American Taxation Association* 25 (Supplement): 33–50.

AICPA. 2000. *Statements on Standards for Tax Services*. New York, NY: AICPA.

Bandy, D. 1996. Limiting tax practice liability. *The CPA Journal* 66 (5): 46–49.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51: 1173–1182.

Beach, L., and T. Mitchell. 1978. A contingency model for the selection of decision strategies. *Academy of Management Review* 3 (July): 439–49.

Bell, T., F. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations Through a Strategic-Systems Lens*. KPMG Peat Marwick.

Biggs, S., T. Mock, and P. Watkins. 1988. Auditors' use of analytical review in audit program design. *The Accounting Review* 63 (January): 148–161.

Bonner, S., J. Davis, and B. Jackson. 1992. Expertise in corporate tax consulting: The issue identification stage. *Journal of Accounting Research* (Supplement): 1–28.

Causey, D., and B. McNair. 1990. *The Tax Professional*. Mississippi State, MS: Mississippi State Press.

Cloyd, C. B. 1995a. The effects of financial accounting conformity on recommendations of tax preparers. *The Journal of the American Taxation Association* 17 (Fall): 50–70.

———. 1995b. Prior knowledge, information search behaviors and performance in tax-research tasks. *The Journal of the American Taxation Association* 17 (Supplement): 62–77.

———. 1997. Performance in tax-research tasks: The joint effects of knowledge and accountability. *The Accounting Review* 72 (January): 111–31.

Collins, K., C. O'Neil, and J. Cathey. 1990. Tax professionals' reactions to return preparer penalties. In *Advances in Taxation*, edited by XXX, 15–43. City, STATE: JAI Press.

Davis, J., and I. Solomon. 1989. Experience, expertise and expert performance in public accounting. *Journal of Accounting Literature*: 150–164.

———, and D. Mason. 2003. Similarity and precedent in tax authority judgment. *The Journal of the American Taxation Association* 25 (Spring): 53–71.

Ferguson, C. 1996. Tax professionals face heightened litigation and traps for the unwary. *The Tax Advisor* 27 (8): 501–502.

Jones, S. M., ed. 1994. *Price Waterhouse Case Studies in Tax*. New York, NY: Price Waterhouse.

Kadous, K., and A. Magro. 2001. The effects of practice risk on the judgments and recommendations of tax professionals. *Contemporary Accounting Research* 18 (3): 441–476.

Kahan, S. 1999. New, needed protection. *The Practical Accountant* 32 (7): 69–72.

Karlin, B. 2003. *Tax Research.* Upper Saddle River, NJ: Prentice Hall.

Keppel, G., and T. Wickens. 2004. *Design and Analysis: A Researcher's Handbook.* Upper Saddle River, NJ: Pearson Prentice Hall.

Khattree, R., and D. Naik. 1999. *Applied Multivariate Statistics With SAS Software.* Cary, NC: SAS Institute.

Klein, G., J. Orasanu, R. Calderwood, and C. Zsambok, eds. 1993. *Decision Making in Action: Models and Methods.* Norwood, NJ: Ablex.

Libby, R., and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society* 18 (5): 425–450.

———. 1995. The role of knowledge and memory in audit judgment. In *Judgment and Decision Making Research in Accounting and Auditing*, edited by R. Ashton, and A. Ashton, 176–206. Boston, MA: Cambridge University Press.

———, and H. Tan. 1997. Tacit managerial versus technical knowledge as determinants of audit expertise in the field. *Journal of Accounting Research* 35 (1): 97–113.

Magro, A. 1999. Contextual features of tax decision-making settings. *The Journal of the American Taxation Association* 21 (Supplement): 63–73.

Payne, J., J. Bettman, and E. Johnson. 1993. *The Adaptive Decision Maker.* Cambridge, U.K.: University of Cambridge Press.

Petty, R., and J. Cacioppo. 1986. The elaboration likelihood model of persuasion. In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, 123–205. City, STATE: Publisher.

Quattrochi, J. 1982. *Federal Tax Research.* New York, NY: Harcourt Brace Jovanovich.

Reder, L. 1987. Strategy selection in question answering. *Cognitive Psychology* 19: 90–138.

Roberts, M., and G. Klersey. 1996. Effects of authoritative guidance and experience on tax decision making. Working paper, The University of Alabama.

Siegler, R. 1988. Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General* 117: 258–275.

Sommerfeld, R., and G. F. Streuling. 1981. *Tax Research Techniques.* Second edition. New York, NY: AICPA.

Spilker, B. 1995. The effect of time pressure and knowledge on key word selection behavior in tax research. *The Accounting Review* 70 (January): 49–70.

Stevens, M. 2000. Managing risk in a tax practice. *The Practical Accountant* 33 (3): 39–41.

Wood, R. 1986. Task complexity: Defining the construct. *Organizational Behavior and Human Decision Processes* 37 (February): 60–82.

THE ACCOUNTING REVIEW
Vol. 80, No. 2
2005
pp. 723–744

The Fleeting Effects of Disclosure Forthcomingness on Management's Reporting Credibility

Molly Mercer
Emory University

ABSTRACT: This study provides a theoretical framework and experimental evidence on how managers' disclosure decisions affect their credibility with investors. I find that in the short-term, more forthcoming disclosure has a positive effect on management's reporting credibility, especially when management is forthcoming about negative news. However, these short-term credibility effects do not persist over time. In the long-term, managers who report positive earnings news are rated as having higher reporting credibility than managers who report negative earnings news, regardless of their previous disclosure decisions.

Keywords: *management credibility; reporting credibility; management disclosure; affect.*

I. INTRODUCTION

Both anecdotal and academic evidence suggest that management's reporting credibility matters. Articles in the popular press often detail managers' credibility struggles and subsequent attempts to regain reporting credibility.[1] The academic literature confirms that managers' credibility concerns are warranted—more credible managers are better able to communicate information to the capital markets (Williams 1996; Mercer 2004). Despite the importance of this topic, there is little systematic evidence on the relation between the communications provided by management and the development of reporting credibility. My paper attempts to fill this void by providing a theoretical framework and

---

[1] Recent examples include articles on AOL Time Warner management's attempt to "restore the company's credibility with investors" (Cullen et al. 2003) and how Honeywell management is "in the credibility penalty box" (Barrett 2003).

This paper is based on my dissertation completed at The University of Texas at Austin. I thank my dissertation chair, Lisa Koonce, for her support and guidance throughout this project, and my other committee members—Keith Brown, Michael Clement, Eric Hirst, and Senyo Tse—for their many helpful suggestions. This paper also benefited greatly from the comments of two anonymous reviewers, Jan Barton, Ananda Ganguly, Kevin Jackson, Steve Kachelmeier, Ron King, Bill Kinney, Jay Koehler, Mary Lea McAnally, Jeff Miller, Mark Peecher, Grace Pownall, Bill Schwartz, Kristy Towry, Rick Tubbs, Greg Waymire, Alex Yen, and seminar participants at The University of Arizona, Emory University, University of Illinois, The University of Iowa, University of Minnesota, Santa Clara University, The University of Texas, University of Utah, Washington University, and the 2002 FARS Midyear Meeting. The Center for Business Measurement and Assurance Services and the Eugene and Dora Bonham Memorial Fund provided financial support for this project.

Editor's note: This paper was accepted by Marlys Gascho Lipe, Editor.

Submitted May 2002
Accepted June 2004

experimental evidence regarding how managers' disclosure decisions affect their credibility with investors.

Specifically, the paper draws on two theories from psychology—attribution theory and affect-based reasoning theory—to develop a model of how the forthcomingness of management's disclosures affects management's reporting credibility (i.e., management's perceived trustworthiness and competence in financial reporting). The model's basic proposition is that management's reporting credibility will be based on different factors in the short- and long-term. In the short-term, reporting credibility will be determined by investors' *cognitive reactions (i.e., thoughts) about the forthcomingness of management's disclosures.* That is, because investors prefer more forthcoming disclosure, greater forthcomingness will increase management's reporting credibility. According to attribution theory, this effect will be especially strong for forthcomingness about negative news.

In the long-term (i.e., as time passes subsequent to managers' disclosure decisions), investors' initial cognitions about these decisions will fade. However, affect-based reasoning theories suggest that investors' affective reactions will remain in memory even after the related cognitions have faded (Damasio 1994; Kida and Smith 1995). Consequently, the model predicts that *affective reactions to the news disclosed* will determine management's reporting credibility in the long-term. That is, managers who previously reported positive news will be perceived as having higher reporting credibility than managers who previously reported negative news, regardless of management's level of forthcomingness about the news.

In short, the model suggests that more forthcoming disclosures, especially those related to negative news, will result in short-term boosts to management's reporting credibility. However, as time passes, investors will no longer reward managers for forthcoming disclosures. Instead, investors' assessments of management's reporting credibility will be based on the firm's financial performance (i.e., whether the firm reported positive or negative news).

This paper also provides experimental tests of the validity of the proposed credibility model. Specifically, I present the results of an experiment in which 244 investors assess management's reporting credibility in various scenarios. The experiment employs a 2 × 2 × 2 between-subjects design that varies (1) management's forthcomingness about unexpected earnings news, (2) whether the news is positive or negative, and (3) whether management's reporting credibility is assessed immediately or two weeks after management's disclosure decision. The experimental results are generally consistent with the model's predictions as described above.

This research is informative on both a theoretical and an applied level. Whereas a useful literature has developed pertaining to the *consequences* of reporting credibility (e.g., Williams 1996; Hirst et al. 1999), little theoretical work has been done on the *determinants* of reporting credibility. Given the significant role that management's reporting credibility plays in a firm's ability to communicate information to investors (Williams 1996; Mercer 2004), developing an understanding of the factors that affect management's reporting credibility is critical. The present study takes an important step in that direction by developing and testing the first theoretically grounded model of how managers' disclosure decisions affect their credibility with investors. This model should be of use to future accounting researchers studying the determinants of reporting credibility, as well as researchers in other fields interested in the development of credibility, such as management, marketing, political science, and sociology.

At an applied level, the results of this study should be informative to firm managers. When trying to build reporting credibility, managers routinely face decisions about the extent to which they should disclose news to investors. My study's results provide guidance to managers regarding the credibility consequences of these decisions. Specifically, they suggest that there are short-term credibility benefits to more forthcoming disclosure, particularly for negative news. This knowledge should be valuable, because reporting credibility is especially important to managers during certain periods of time, such as when they need to raise capital (Healy and Palepu 1995). However, my results also suggest that efforts to increase long-term reporting credibility via disclosure policies may be misplaced, as the credibility consequences of disclosure decisions are fleeting.

The paper proceeds as follows. Section II draws on theories from cognitive and social psychology to derive a model of how the forthcomingness of management's disclosures affects its reporting credibility in the short- and long-term. Section III presents an experiment that tests the model's predictions, and Section IV reports the experiment's results. Section V provides summary remarks and discusses the implications of this research for managers and academics.

II. MODEL DEVELOPMENT

Overview

I define management's reporting credibility as *investors' beliefs about management's trustworthiness and competence in financial disclosure* (Hovland et al. 1953; Giffin 1967).[2] This section draws on several theories from psychology to develop a model of when and how management's disclosure decisions affect its reporting credibility. The model suggests that after receiving company news and inferring management's level of forthcomingness about the news, investors will experience both cognitive and affective reactions. Cognitive reactions take the form of thoughts (Markman and Gentner 2001) about the news and management's forthcomingness, and affective reactions involve positive or negative evaluative reactions (Slovic et al. 2002) to the news and management's forthcomingness.[3]

The model's basic proposition is that the credibility consequences of managers' disclosure decisions will differ in the short- and long-term. Specifically, the model predicts that in the short-term, investors' cognitive reactions to management's disclosure decisions determine the credibility consequences of those decisions. That is, because investors prefer more forthcoming disclosures, they will deem managers who provide such disclosures to be more credible than managers who do not. Moreover, the model predicts that investors will consider the *reasons* for management's disclosure decisions and that these thoughts will influence the credibility effects of disclosure decisions.

The model also predicts that as time passes, investors' cognitive reactions to management's disclosure decisions will fade. However, affective reactions make stronger imprints in memory than cognitive reactions (Damasio 1994; Kida and Smith 1995), so investors will retain their overall affective reaction to an event long after their cognitive reactions have faded. Thus, when investors assess management's reporting credibility in the long-term, their initial overall affective reaction will form the basis for their assessments. Prior literature suggests that investors' overall reactions will be driven primarily by the valence of the news disclosed (i.e., whether the news disclosed was positive or negative) rather than by whether management was forthcoming about the news (Kasznik and Lev 1995). Consequently, the model predicts that, in the long-term, management's reporting credibility will be a function of whether the firm reported positive or negative news, rather than management's forthcomingness about the news. These short- and long-term credibility effects are described in more detail below.

---

[2] The idea that management's reporting credibility represents investors' beliefs about management is consistent with reputation studies in the economics literature. For example, Axelrod (1984) defines reputation as the embodiment of beliefs about the strategy that a particular person will use.

[3] The term "affect" has been used to refer to various constructs, including evaluations, moods, and emotions (Fiske and Taylor 1991). Consistent with recent articles that examine the role of affect in decision making (e.g., Mellers et al. 1997; Finucane et al. 2000), I focus on affect as an evaluation (i.e., as a positive or negative emotional response associated with an event).

Short-Term Credibility Effects

The model's predictions regarding the short-term credibility effects of managers' disclosure decisions center on investors' cognitions about management's forthcomingness. However, before describing these predictions, it is important to consider the factors that influence investors' perceptions of disclosure forthcomingness. Three facets of a firm's disclosures appear to be important predictors of perceived disclosure forthcomingness—the accuracy, completeness, and timeliness of a firm's disclosures. Disclosure accuracy is the most obvious factor in the perceived forthcomingness of management disclosures. *Ceteris paribus*, disclosures that are revealed to be inaccurate are more likely to be perceived as nonforthcoming than more accurate disclosures (e.g., Barrett 2003; Rama 2003). However, the completeness and timeliness of management's disclosures also influence perceptions of forthcomingness. For example, a company's annual report might report accurate overall financial results but not disclose enough detail for investors to determine the performance of certain critical business segments. Investors may regard such incomplete disclosures as nonforthcoming despite their accuracy (e.g., Silverman 2002; Steinberg 2003). Investors also may regard untimely disclosures as nonforthcoming. For example, firms that fail to provide earnings warnings are often perceived as less forthcoming than firms that provide such disclosures (e.g., Elgin 2003; Olver 2003).

The few academic studies that consider the relation between disclosure forthcomingness and management's reporting credibility focus on the *short-term* credibility effects of forthcoming disclosures (Libby and Tan 1999; Tan et al. 2002). Both Libby and Tan (1999) and Tan et al. (2002) ask investors to rate management's credibility immediately after receiving information about management's forthcomingness. These studies show that investors rate managers who provide forthcoming disclosures as more credible than managers whose disclosures are less forthcoming. Specifically, Libby and Tan (1999) find that managers who provide warnings about unexpected earnings are judged to have higher integrity than managers who do not provide such disclosures. Tan et al. (2002) manipulate the forthcomingness of management's disclosures by varying disclosure accuracy rather than disclosure timeliness and find similar results. That is, managers who accurately forecast earnings are rated as more competent and trustworthy than managers who significantly overstate or understate earnings. In short, these studies find that forthcoming managers have higher credibility than nonforthcoming managers, suggesting that disclosure forthcomingness affects management's reporting credibility, at least in the short-term. This leads to the following hypothesis:

> **H1a:** In the short-term, forthcoming disclosure will increase management's reporting credibility and a lack of forthcomingness will decrease management's reporting credibility.

Attribution theory from social psychology provides a basis for refining H1a. Attribution theory describes how people impute the causes of others' behaviors (Fiske and Taylor 1991). A key tenet of attribution theory is that when situational explanations for an actor's behavior are not apparent, people are more likely to attribute the behavior to an enduring dispositional trait of the actor (Jones and Davis 1965). When situational forces are apparent, people are less likely to make a dispositional attribution. In the management disclosure context, attribution theory predicts that disclosures that are at odds with managers' personal incentives are more likely to be attributed to a dispositional characteristic (e.g., the manager's honesty or trustworthiness) than disclosures that are consistent with the managers' incentives. In other words, because managers tend to have fewer personal incentives to disclose negative news than positive news (Dimma 1996; Hughes 1997), investors are more likely to attribute forthcoming negative news disclosures to management's honesty and trustworthiness than forthcoming positive news disclosures.

Research in attribution theory also suggests that investors will spend more time thinking about negative outcomes (Weiner 1985), and thus are more likely to think about the reasons for negative news disclosure decisions. Consequently, investors are more likely to think about management's role in negative news disclosure decisions than positive news disclosure decisions (Ybarra 2002). Such additional thought is likely to (1) heighten the favorable dispositional attributions investors may already be inclined to make regarding managers who are forthcoming about negative news, and (2) heighten the unfavorable dispositional attributions investors may be inclined to make regarding managers who are not forthcoming about negative news. With these principles in mind, I expect that managers will receive more credit for forthcoming disclosure and more blame for nonforthcoming disclosure when the disclosure pertains to negative news. This reasoning leads to Hypothesis 1b:

> **H1b:** In the short-term, forthcoming disclosure of negative news will increase management's reporting credibility more than forthcoming disclosure of positive news. A lack of forthcomingness about negative news will decrease management's reporting credibility more than a lack of forthcomingness about positive news.

A graphical summary of the short-term credibility predictions is provided in Panel A of Figure 1.

Long-Term Credibility Effects

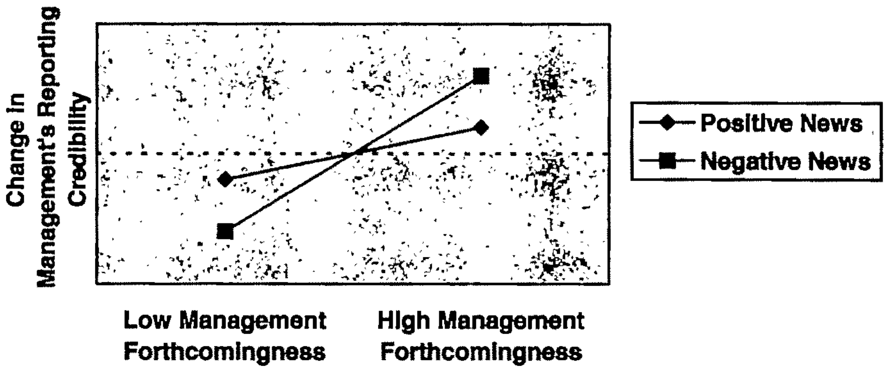
Central to my model is the claim that the short-term credibility effects predicted in H1a and H1b—effects that are driven by investors' cognitive reactions to management's disclosure decisions—may not hold over time. Rather, in the long-term, perceptions of management's reporting credibility will be driven primarily by investors' affective reactions to the news disclosed. In making this prediction, I draw on an affect-based model of financial decision-making proposed by Kida and Smith (1995).

Kida and Smith (1995) argue that when investors experience an event, both the event and the overall affective reaction to that event are encoded in memory. Further, they posit that affective reactions create stronger imprints in memory than the specific events underlying those reactions. As time passes subsequent to a disclosure decision, investors' memories of the detailed information underlying their credibility assessments will fade. However, Kida and Smith's (1995) research suggests that investors' affective reactions will be accessible long after the specific details underlying them have faded. Consequently, when investors assess management's reporting credibility in the long-term, they will reconstruct the details of managers' disclosure decisions to be consistent with their initial overall affective reaction (Clore et al. 1994; Kida and Smith 1995). Thus, investors' long-term credibility assessments will be driven by their initial overall affective reaction.

This argument suggests that the long-term credibility consequences of managers' disclosure decisions will depend on the source of investors' overall affective reactions. After receiving news and inferring the management's forthcomingness about the news, investors

FIGURE 1
Predicted Changes in Management's Reporting Credibility as a Function of News Valence and Management's Forthcomingness

Panel A: Short-Term Predictions



Panel B: Long-Term Predictions

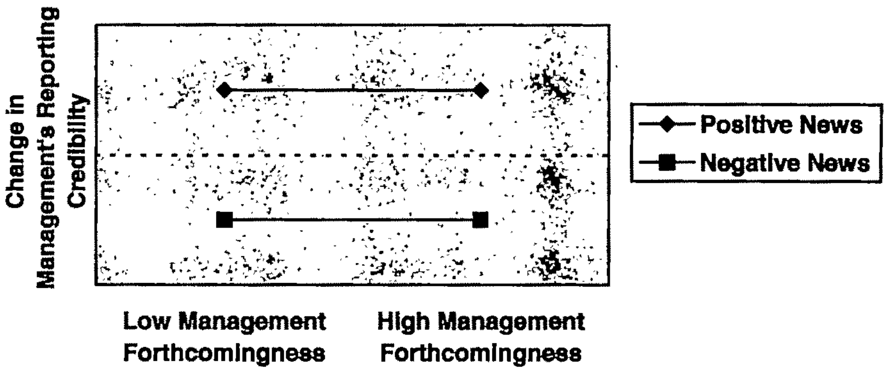


Figure 1 summarizes the predicted short-term (Panel A) and long-term (Panel B) change in management's reporting credibility as a function of news valence and management's forthcomingness. Values above (below) the dotted line represent predicted increases (decreases) in management's reporting credibility.

will experience affective reactions to (1) the valence of the news itself (positive or negative), and (2) management's forthcomingness about the news (forthcoming or not forthcoming). Positive news will result in a positive affective reaction, while negative news will result in a negative affective reaction. Similarly, forthcoming disclosure will result in positive affect, and a lack of forthcomingness will result in negative affect. In some cases, these affective reactions will be in conflict, such as when there is negative news (which results in negative affect) but management is forthcoming about the news (which results in positive affect).

Investors' *overall* affective reaction to these events will depend on which of these two affective reactions is stronger (Damasio 1994). A paper by Kasznik and Lev (1995) provides reason to believe that investors' affective reactions to news valence will be stronger than their affective reactions to management's forthcomingness.[4] Specifically, Kasznik and Lev (1995) find that earnings news is a stronger predictor of stock market reactions than whether management was forthcoming about the news (i.e., whether management provides an earnings warning). Consequently, I expect that positive (negative) news will result in positive (negative) overall affect being encoded in memory, regardless of management's forthcomingness about the news.

If investors' overall affective reactions are the primary driver of their long-term credibility judgments and if news valence determines these affective reactions, then the long-term credibility consequences of managers' disclosure decisions will depend on the valence of the news disclosed. As a result, more forthcoming disclosures will not necessarily result in greater reporting credibility in the long-term. Rather, managers of firms that report positive news will have greater reporting credibility than managers of firms that report negative news, regardless of the managers' forthcomingness.

To summarize, as time passes subsequent to managers' disclosure decisions, investors' cognitions about these decisions will fade. Consequently, investors who assess management's reporting credibility at a later date will base their assessments on their initial overall affective reaction rather than these cognitions. The model also predicts that an investor's overall affective reaction will be driven by the valence of the news (positive or negative) rather than management's forthcomingness about the news. As a result, in the long-term, investors' perceptions of management's reporting credibility will be based on whether the firm previously reported positive or negative news. That is, managers who reported positive news will be rewarded with increased credibility and managers who reported negative news will be penalized with decreased credibility. In sharp contrast to the model's short-term predictions, management's forthcomingness is not expected to be the primary driver of its reporting credibility in the long-term. This reasoning is summarized in Hypothesis 2 and can be seen graphically by comparing Panels A and B of Figure 1.

> **H2:** In the long-term, managers who previously reported positive news will experience increases in credibility and managers who previously reported negative news will experience decreases in credibility, regardless of their perceived forthcomingness.

III. EXPERIMENT

Participants

To test my hypotheses, I conducted an experiment with 244 M.B.A. students at a large state university. Participants were recruited via an email notice and were paid a flat wage of $10 plus the chance to win additional money in a random lottery. On average, study participants had six years of work experience and had completed three accounting and two finance classes; most (83 percent) had previously invested in individual common stocks. I used M.B.A. students as participants in this study because their business experience, investing experience, and coursework should provide them with a basis for understanding managers' disclosure incentives (Elliott et al. 2004). Previous research indicates that some understanding of the incentives managers have to disclose or withhold information is necessary for the presentation of the hypothesized short-term credibility effects (Robertson and Rossiter 1974). For example, if participants do not understand that managers have incentives to withhold information, then they will not reward forthcoming managers with short-term increases in reporting credibility.

---

[4] If this prediction is incorrect, and affective reactions to management's forthcomingness are the primary determinant of investors' overall affective reactions, then H2 will not hold. In this case, management's forthcomingness, rather than news valence, will drive reporting credibility in the long-term.

Experimental Design and Task

The experiment uses a 2 × 2 × 2 between-subjects design that varies the following three variables: management's forthcomingness about unexpected earnings news (Management Forthcomingness: Low, High), the valence of the earnings news (News Valence: Positive, Negative), and the amount of time between reception of the earnings news and participants' assessments of management's reporting credibility (Time: Short-term, Long-term). An experimental approach to investigations of the credibility effects of disclosure forthcomingness has several advantages. First, using an experiment allows me to hold constant factors other than management's forthcomingness. Such control is important when studying the determinants of reporting credibility because firms with highly credible managers are fundamentally different in terms of earnings and returns than those with less credible managers (Lang and Lundholm 1993). Second, experimentation allows for the collection of data on the credibility effects of specific disclosure decisions. Archival researchers can use financial analysts' ratings of a firm's disclosures (AIMR reports) to measure managers' reporting credibility. However, AIMR reports are annual ratings, and the time lag between a particular management disclosure and the subsequent AIMR report makes it difficult to use these reports to infer the credibility effects of specific disclosures.

Experimental participants assumed the role of investment club members who had recently purchased common stock in Dentex, a dental supply company. Each participant was given background information about Dentex, including excerpts from management's discussion and analysis and historical financial statement data. Participants also were informed that the consensus analyst forecast for Dentex's upcoming quarterly earnings was $0.52. After reviewing this information, participants provided pre-test assessments of management's reporting credibility.

Participants in the High Management Forthcomingness condition then received a voluntary disclosure from management stating that actual quarterly earnings were expected to differ from the consensus forecast. Specifically, participants in the Positive (Negative) News Valence condition received the following management disclosure.

> On March 15, Dentex management voluntarily issued the following press release:
>
> In a presentation to industry analysts today, Dentex, Inc. CEO Murray Levine indicated that the company currently expects earnings per share will be $0.60 ($0.44) for the quarter ending March 31. This estimate is $0.08 above (below) the current consensus analyst forecast of $0.52.[5]

Participants in the Low Management Forthcomingness condition did not receive this disclosure.

Subsequently, all participants answered a number of demographic questions regarding their investing experience, work experience, and prior coursework. After providing these assessments, participants received information about the firm's actual quarterly earnings.

---

[5] Brown (2001) finds that approximately 75 percent of profit firms and 43 percent of loss firms report earnings that are less than $0.05 from the consensus analyst forecast. Thus, the amount of unexpected earnings in this study ($0.08) is larger than average. Relatively large unexpected earnings should enhance differential perceptions of management's forthcomingness across the High Forthcomingness and Low Forthcomingness conditions (Tan et al. 2002; Libby et al. 2002, 798). Smaller unexpected earnings may yield smaller effect sizes but should show the same overall pattern of responses.

Actual earnings were $0.60 in the Positive News Valence condition and $0.44 in the Negative News Valence condition. Participants then answered several manipulation check and affective reaction questions.

Participants in the Short-term condition answered additional questions immediately after viewing actual earnings. Specifically, they re-assessed management's reporting credibility, answered questions about their cognitive processing, and indicated their willingness to rely on a subsequent forecast issued by management. Long-term participants answered these same questions two weeks after completing the experimental materials. I chose a two-week time delay because two weeks is a sufficient amount of time for memory reconstruction to commence (Spiro 1980; Kida et al. 1998), but not so long that it threatens the participant response rate.[6]

Dependent Variables

Management's Reporting Credibility

I drew six questions from two widely accepted source credibility scales—McCroskey (1966) and Leathers (1992)—to measure management's reporting credibility (see Appendix). Three questions related to management's competence in financial reporting. These questions asked investors to assess management's competence, knowledge, and qualifications for providing financial disclosures. Three additional questions related to management's trustworthiness in financial reporting. These questions focused on investors' perceptions of management's trustworthiness, honesty, and truthfulness in its financial disclosures.

To ensure that these questions capture one underlying credibility construct, I conducted a reliability analysis on participants' responses. The results of this analysis yielded a Cronbach's alpha of 0.78, suggesting that the scale is reliable (Nunnally 1978, 245). Consequently, I formed a composite measure of management's reporting credibility by summing participants' responses to the six questions.[7] Because the hypotheses are framed in terms of *changes* in management's reporting credibility, I then calculated a credibility difference score for each participant (i.e., post-test composite reporting credibility minus pre-test composite reporting credibility).

As noted previously, prior research has demonstrated the important consequences of management's reporting credibility. These studies show that investors are more likely to rely on the disclosures of highly credible managers, making it easier for these managers to communicate information to the capital markets (Williams 1996; Hirst et al. 1999). The present paper examines the *determinants* of management's reporting credibility, with an emphasis on how the forthcomingness of management disclosures affects reporting credibility. Consequently, my primary dependent variable is changes in management's reporting credibility. However, to ensure that participants' ratings of reporting credibility have significant consequences, I also examine whether changes in perceived reporting credibility affect participants' willingness to rely on subsequent management disclosures. Specifically, participants were shown a management earnings forecast for a subsequent quarter and asked to express their willingness to rely on this forecast when predicting earnings. A positive news earnings forecast was chosen for this task because prior research shows that reporting credibility is especially important for increasing the believability of positive news disclosures (Williams 1996).

---

[6] Although I only use Long-term participants' two-week-later responses, I asked all participants to return after two weeks to prevent them from guessing the manipulations. To motivate participants to return, those who completed both parts of the study were entered into a $1500 lottery ($1000 first prize and $500 second prize). Approximately 89 percent of participants returned after two weeks.

[7] Inferentially identical results appear when management's reporting credibility is measured with credibility factor scores derived by conducting a factor analysis on participants' responses to the six credibility questions.

Process Measures

My model also identifies the mechanisms through which managers' disclosure decisions affect reporting credibility. The model proposes that the short-term credibility consequences of managers' disclosure decisions are driven primarily by investors' cognitive reactions—in particular, the attributions they make for managers' disclosure decisions. To measure participants' attributions, I provided a list of six potential explanations for management's disclosure decision: pressure from analysts and investors, self-interest, management's honesty, predictability of the firm's earnings, a desire for the firm's stock price to be consistent with economic reality, and legal liability concerns. Participants ranked these explanations from 1 to 6, where 1 (6) was the factor that they believed was the most (least) important cause of management's disclosure decision. Because my theoretical framework suggests that *dispositional* attributions lead to changes in management's perceived credibility, I am primarily interested in attributions to "management's honesty/dishonesty." I use the relative rank of this variable as my primary measure of investors' cognitive reactions.

The model predicts that investors' long-term credibility judgments are driven by their affective reactions to the news valence, rather than their cognitive reactions to management's disclosure decisions. Consistent with prior research on the role of affect in decision-making (Mellers et al. 1997; Kida et al. 1998), I measured participants' affective reactions to the news valence by asking them to indicate their agreement with the following statement, "The unexpected earnings caused me to feel good." Participants indicated their agreement on a seven-point Likert-type scale, with endpoints 1 = Strongly Disagree and 7 = Strongly Agree. Responses to this question capture both the direction and strength of participants' affect. That is, responses above (below) the scale midpoint indicate positive (negative) affect; the higher (lower) the number, the more positive (negative) the affective reaction.[8]

IV. RESULTS

Manipulation Checks

Responses to the manipulation check questions indicate that all manipulations were successful. Among participants in the High (Low) Management Forthcomingness condition, 98 percent (98 percent) correctly indicated that they did (did not) receive a warning about the company's unexpected earnings. Among participants in the Positive (Negative) News Valence condition, 100 percent (99 percent) correctly indicated the direction of the unexpected earnings news. Excluding participants who failed one or more of the manipulation checks does not materially change any of the results reported below.

Effects of Management's Disclosure Decisions on Its Reporting Credibility

The proposed model makes a number of specific predictions about the short- and long-term credibility effects of managers' disclosure decisions. Underlying these predictions is the assumption that the credibility consequences of disclosure decisions differ in the short- and long-term. Thus, before testing the model's specific predictions, I test this general assumption. Specifically, I conduct a three-way ANOVA with Management Forthcomingness, News Valence, and Time as independent variables and participants' credibility difference scores as the dependent variable. As predicted, the ANOVA reveals a significant three-way interaction term ($F = 4.14$, two-tailed $p = 0.04$). This result suggests that the credibility effects of more forthcoming disclosure differ in the short- and long-term. I discuss these short- and long-term effects in more detail in the following two sections.

---

[8] Participants also evaluated the statement, "The unexpected earnings caused me to feel bad" on a similar Likert-type scale (1 = Strongly Disagree and 7 = Strongly Agree). Participants' responses to this "feel bad" question are inversely related to their responses to the "feel good" question, suggesting that the "feel good" question is able to capture both positive and negative affect.

Short-Term Credibility Effects

Recall that the model offers two hypotheses regarding the short-term credibility effects of management's disclosure decisions: Hypothesis 1a predicts that forthcoming disclosures will increase management's reporting credibility and a failure to provide forthcoming disclosures will decrease management's reporting credibility. Hypothesis 1b predicts that these effects will be more pronounced for negative news than for positive news. To test these hypotheses, I first conduct an omnibus two-way ANOVA with Management Forthcomingness and News Valence as independent variables and Short-term participants' credibility difference scores as the dependent variable (see Table 1). The ANOVA shows a statistically significant main effect for Management Forthcomingness ($F = 52.08$, two-tailed $p < 0.01$) and a statistically significant Management Forthcomingness × News Valence interaction ($F = 5.17$, two-tailed $p = 0.03$). These results suggest that decisions about whether to provide forthcoming disclosures influence management's reporting credibility and that the magnitude of this influence differs for positive and negative news.

If forthcoming disclosure *increases* reporting credibility and failure to provide forthcoming disclosure *decreases* reporting credibility as proposed in H1a, then credibility difference scores for participants in the High (Low) Management Forthcomingness conditions should be positive (negative). In support of H1a, credibility difference scores in the High Management Forthcomingness conditions are greater than zero (mean = 2.63, $t = 6.07$, one-tailed $p < 0.01$), and credibility difference scores in the Low Management Forthcomingness conditions are less than zero (mean = −2.73, $t = 4.43$, one-tailed $p < 0.01$). These relations hold across both Positive and Negative levels of News Valence. That is, credibility difference scores in the High Management Forthcomingness condition are greater than zero for both Positive News (mean = 2.15, $t = 3.76$, one-tailed $p < 0.01$) and Negative News (mean = 3.13, $t = 4.82$, one-tailed $p < 0.01$), whereas credibility difference scores in the Low Management Forthcomingness condition are less than zero for both Positive News (mean = −1.52, $t = 2.77$, one-tailed $p < 0.01$) and Negative News (mean = −3.94, $t = 3.68$, one-tailed $p < 0.01$).

Hypothesis 1b predicts that the effects of management's forthcomingness on its reporting credibility will be more pronounced for negative news. The data provide partial support for H1b. Participants' credibility difference scores decreased more in the Low Management Forthcomingness/Negative News condition (mean = −3.94) than in the Low Management Forthcomingness/Positive News condition (mean = −1.52) ($t = 2.01$, one-tailed $p = 0.03$).[9] Further, participants' credibility differences scores increased more in the High Management Forthcomingness/Negative News condition (mean = 3.13) than in

---

[9] Tables 1 and 2 show unequal variances across the experimental conditions. I therefore use t-tests that do not assume equal variances when analyzing differences across conditions.

TABLE 1

Short-Term Effects of News Valence and Management's Forthcomingness on Management's Reporting Credibility

Panel A: Mean (Standard Deviation) Changes in Management's Reporting Credibility

	Positive News	**Negative News**	
Low Management Forthcomingness	−1.52 (3.11) n = 32	−3.94 (6.06) n = 32	−2.73 (4.94) n = 64
High Management Forthcomingness	2.15 (3.24) n = 32	3.13 (3.56) n = 30	2.63 (3.41) n = 62
	0.32 (3.66) n = 64	−0.52 (6.12) n = 62	

Panel B: ANOVA Results

	df	**MSE**	**F-statistic**	**p-value**
Management Forthcomingness	1	911.28	52.08	.00
News Valence	1	16.43	0.94	.34
Management Forthcomingness × News Valence	1	90.57	5.17	.03

Table 1 shows the short-term effects of news valence and management's forthcomingness on management's reporting credibility. Prior to viewing the experimental manipulations, participants provided pre-test assessments of management's reporting credibility by answering six questions (see Appendix). A composite measure of management's reporting credibility was computed by summing participants' responses to these questions. Participants in the High Management Forthcomingness condition were then provided with a voluntary disclosure from management stating that actual quarterly earnings were expected to differ from the consensus analyst forecast. Specifically, participants in the Positive (Negative) News condition were told that earnings per share was expected to be $0.08 above (below) the consensus analyst forecast. Participants in the Low Management Forthcomingness condition did not receive this disclosure. Subsequently, all participants were provided with the firm's actual quarterly earnings, which were $0.08 above (below) the consensus forecast for Positive (Negative) News participants. Immediately after viewing these materials, participants provided post-test assessments of management's reporting credibility by answering the same credibility questions used in the pre-test. Panel A reports the change in management's reporting credibility (computed as post-test composite reporting credibility minus pre-test composite reporting credibility) by experimental condition. Panel B reports an ANOVA that tests the short-term effects of news valence and management's forthcomingness on changes in management's reporting credibility.

the High Management Forthcomingness/Positive News condition (mean = 2.15), though this difference did not reach statistical significance (t = 1.12, one-tailed p = 0.13).[10]

To summarize, I find that in the short-term, management's decisions about whether to provide forthcoming disclosures affect its reporting credibility. Forthcoming disclosure results in increased reporting credibility and a lack of forthcoming disclosure results in decreased reporting credibility. Negative news disclosure decisions appear to be especially important in determining reporting credibility—managers who are not forthcoming about negative news experience greater credibility losses than managers who are not forthcoming about positive news. A graphical summary of these effects appears in Panel A of Figure 2.

---

[10] If participants were more surprised by the negative earnings news than the positive earnings news, then it is possible that the *expectedness* of the news, rather than the news valence, per se, drove their responses. To test for this possibility, I asked participants to evaluate the expectedness of the earnings information. The results do not indicate that participants in the Negative News condition were significantly more surprised by the unexpected earnings than participants in the Positive News condition (t = 0.47, two-tailed p = 0.64). Therefore, the expectedness of the news cannot explain the observed effects of news valence.

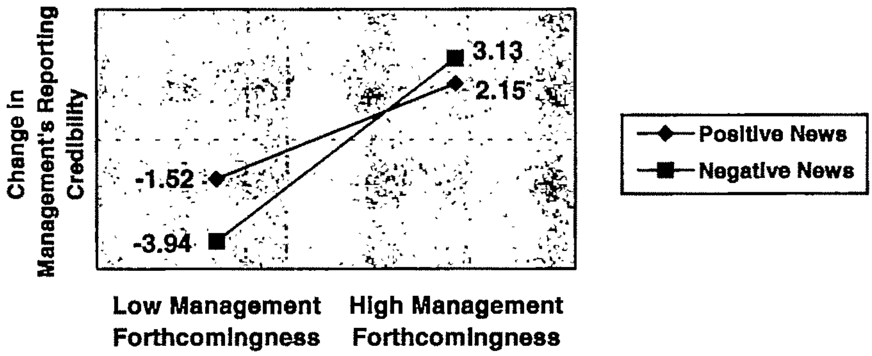
Long-Term Credibility Effects

The model also predicts that the credibility consequences of disclosure forthcomingness will be short-lived. Specifically, the model predicts that news valence (i.e., whether the firm reports positive or negative news), rather than management's forthcomingness, will drive

FIGURE 2

Changes in Management's Reporting Credibility as a Function of News Valence and Management's Forthcomingness

Panel A: Changes in Management's Reporting Credibility in the Short-Term



Panel B: Changes in Management's Reporting Credibility in the Long-Term

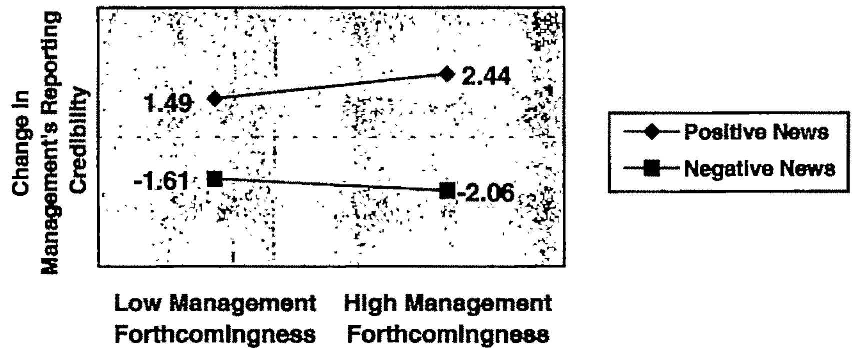


Figure 2 reports the mean short-term (Panel A) and long-term (Panel B) change in management's reporting credibility by experimental condition. Prior to exposure to the experimental manipulations (see Table 1 for descriptions), participants provided pre-test assessments of management's reporting credibility by answering six credibility questions (see Appendix). A composite measure of management's reporting credibility was computed by summing each participant's responses to these questions. Participants in the Short-term conditions answered the questions a second time immediately following exposure to the experimental manipulations, while participants in the Long-term conditions provided post-test assessments two weeks after exposure to the experimental manipulations. Credibility difference scores were computed by subtracting participants' pre-test composite credibility assessments from their post-test composite credibility assessments. Positive (negative) credibility difference scores indicate increases (decreases) in management's reporting credibility.

management's reporting credibility in the long-term. Consequently, H2 predicts that managers who report positive news will experience long-term credibility gains and managers who report negative news will experience long-term credibility losses, regardless of their forthcomingness about the news. In statistical terms, this suggests that a two-way ANOVA on Long-term participants' credibility difference scores will yield a statistically significant main effect for News Valence only. The data support this prediction (see Table 2). The ANOVA yields a significant main effect for News Valence ($F = 16.77$, two-tailed $p < 0.01$) and insignificant effects for Management Forthcomingness ($F = 0.07$, two-tailed $p = 0.79$) and the Management Forthcomingness × News Valence interaction term ($F = 0.57$, two-tailed $p = 0.45$). These results suggest that News Valence is the primary predictor of participants' beliefs about management's reporting credibility in the long-term.

If, as suggested in H2, reporting positive news leads to long-term credibility *increases* and reporting negative news leads to long-term credibility *decreases*, then the credibility difference scores for participants in the Positive (Negative) News conditions should be positive (negative). The data support H2 across both High and Low levels of Management Forthcomingness. In the long-term, Positive News participants' credibility difference scores are greater than zero, regardless of whether management was forthcoming (mean = 2.44, $t = 3.11$, one-tailed $p < 0.01$) or not forthcoming (mean = 1.49, $t = 2.89$, one-tailed $p < 0.01$) about the positive news. This result suggests that positive news has positive effects

TABLE 2
Long-Term Effects of News Valence and Management's Forthcomingness on Management's Reporting Credibility

Panel A: Mean (Standard Deviation) Changes in Management's Reporting Credibility

	Positive News	**Negative News**	
Low Management Forthcomingness	1.49 (2.53) n = 24	−1.61 (4.89) n = 29	−0.21 (4.26) n = 53
High Management Forthcomingness	2.44 (4.16) n = 28	−2.06 (6.57) n = 26	0.27 (5.86) n = 54
	2.00 (3.50) n = 52	−1.83 (5.70) n = 55	

Panel B: ANOVA Results

	df	**MSE**	**F-statistic**	**p-value**
Management Forthcomingness	1	1.69	0.07	.79
News Valence	1	385.23	16.77	.00
Management Forthcomingness × News Valence	1	13.02	0.57	.45

Table 2 presents the long-term effects of news valence and management's forthcomingness on management's reporting credibility. Participants in the Long-term conditions made post-test assessments of management's reporting credibility two weeks after viewing the experimental manipulations. See Table 1 for descriptions of the manipulations. Panel A reports the change in management's reporting credibility (computed as post-test composite reporting credibility minus pre-test composite reporting credibility) by experimental condition. Panel B reports an ANOVA that tests the long-term effects of news valence and management's forthcomingness on changes in management's reporting credibility.

on management's reporting credibility—even when management is not forthcoming about the news. Perhaps more importantly, the credibility difference scores of participants in the Positive News/High Management Forthcomingness condition are not significantly greater than those in the Positive News/Low Management Forthcomingness condition ($t = 1.01$, two-tailed $p = 0.32$). This result suggests that there are few, if any, long-term credibility consequences to withholding positive news.

Negative News participants report credibility difference scores that are less than zero both when management is forthcoming (mean $= -2.06$, $t = 1.60$, one-tailed $p = 0.06$) and when management is not forthcoming (mean $= -1.61$, $t = 1.78$, one-tailed $p = 0.04$). In other words, managers who ultimately report negative news lose reporting credibility, regardless of their forthcomingness. In addition, the credibility difference scores of participants in the Negative News/High Management Forthcomingness condition are not significantly greater than those of participants in the Negative News/Low Management Forthcomingness condition ($t = 0.28$, two-tailed $p = 0.78$). Thus, it appears that there are no long-term credibility benefits to being forthcoming about negative news.

In sum, I find that in the long-term, management's reporting credibility is unrelated to the forthcomingness of its disclosures. Instead, management's reporting credibility is driven by the valence of the news that is ultimately reported. Specifically, managers who report negative news experience decreased reporting credibility, and managers who report positive news experience increased reporting credibility, regardless of their forthcomingness about the news. Panel B of Figure 2 summarizes these long-term credibility effects.

Understanding the Process

Overview

The model not only predicts that the credibility consequences of managers' disclosure decisions will differ in the short- and long-term, but it also identifies factors that may explain why these effects occur. Specifically, the model posits that investors' *cognitive reactions* to disclosure decisions will drive their perceptions of management's reporting credibility in the short-term, but that investors' *affective reactions* to the news valence will drive their perceptions of management's reporting credibility in the long-term. I test these hypotheses about the determinants of management's reporting credibility below. In addition, I explore the potential consequences of changes in management's reporting credibility. Specifically, I examine whether changes in management's reporting credibility are reflected in investors' willingness to rely on management's subsequent disclosures.

I use the AMOS path analysis program (Arbuckle 1999) on data collected in my experiment to simultaneously test these determinants and consequences. Path analysis extends the regression model by allowing for simultaneous examination of both direct and indirect effects among a set of variables (Kline 1998). This technique is used to assess the overall fit of a causal model as well as the significance of specific links within the model. In the present context, path analysis provides a parsimonious and powerful method for identifying the roles that investors' cognitive and affective reactions play in their beliefs about management's reporting credibility in the short- and long-term, and whether changes in management's reporting credibility, in turn, affect investors' willingness to rely on management's subsequent disclosures.

Determinants of Reporting Credibility

The model predicts that investors will attribute forthcoming management disclosures to management's dispositions, and that these cognitions will lead to short-term increases in management's reporting credibility. However, because cognitions fade from memory over time, the model offers a different prediction about the determinants of management's reporting credibility in the long-term. Investors' affective reactions, which are driven largely by news valence, make stronger imprints in memory than their cognitive reactions. Consequently, the model also predicts that investors' affective reactions to the company's news—rather than their cognitive reactions to management's forthcomingness about the news—will drive investors' beliefs about management's reporting credibility in the long-term.

I test these predictions by conducting two path analyses (see Figure 3). These analyses test the effects of management's forthcomingness about positive and negative news on participants' cognitive and affective reactions, and the effects of these reactions, in turn, on participants' credibility difference scores. The first path analysis examines Short-term participants' responses (Panel A), and the second path analysis examines Long-term participants' responses (Panel B).

The data appear to fit the model well. Comparative fit indices (CFIs) greater than 0.90 suggest good model fit (Kline 1998, 131), and the Short- and Long-term path models have CFIs of 0.99 and 0.98, respectively. Further, nested models confirm that the overall path is different in the Short- and Long-term ($\chi^2 = 5.78$, two-tailed $p = 0.06$).

Results for the individual model links are also as predicted. Panel A of Figure 3 shows a statistically significant link between Short-term participants' cognitive reactions and their credibility difference scores ($\beta = 0.25$, two-tailed $p < 0.01$). Thus, as expected, cognitive reactions are a significant predictor of management's reporting credibility in the short-term. Also as expected, the link between Short-term participants' affective reactions and their credibility difference scores is not significant ($\beta = 0.12$, two-tailed $p > 0.10$). This result suggests that affective reactions are not a significant predictor of management's reporting credibility in the short-term.

The model predicts that affective reactions will be the primary determinant of participants' credibility assessments in the long-term. In support of this hypothesis, Panel B of Figure 3 shows a statistically significant link between Long-term participants' affective reactions and their credibility difference scores ($\beta = 0.27$, two-tailed $p < 0.01$). The link between Long-term participants' cognitive reactions and their credibility difference scores is not statistically significant ($\beta = -0.08$, two-tailed $p > 0.10$). These results indicate that affective reactions—rather than cognitive reactions—are the primary driver of Long-term participants' beliefs about management's reporting credibility.

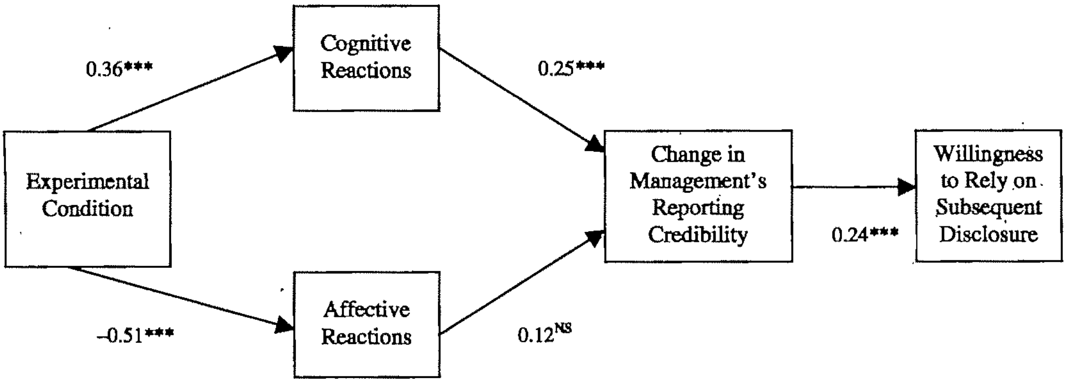
Taken together, the path analyses provide strong support for the model. Investors' cognitive reactions to management's disclosure decisions influence investors' beliefs about management's reporting credibility in the short-term, but not in the long-term. In contrast, investors' affective reactions influence their beliefs about management's reporting credibility in the long-term, but play little or no role in the short-term. Thus, the path analyses support the model's contention that managers' disclosure decisions have different credibility consequences in the short- and long-term because investors' credibility assessments are informed by different factors over time.

Consequences of Reporting Credibility

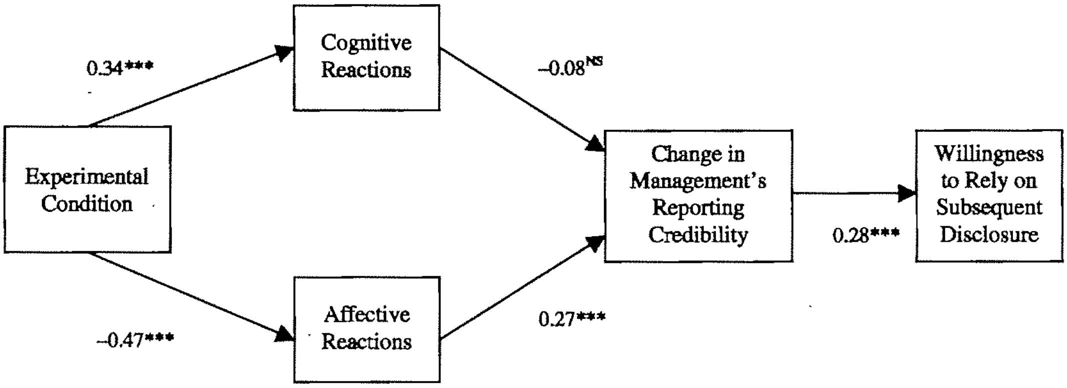
The path analyses also examine the consequences of gains or losses to management's reporting credibility. Prior studies suggest that management's reporting credibility influences its ability to communicate information to investors. Specifically, investors are more likely to rely on disclosures from highly credible managers (Williams 1996; Hirst et al. 1999). To examine whether the credibility effects demonstrated in this study are consequential, all

FIGURE 3
Path Analysis Results—Determinants and Consequences of Management's Reporting Credibility

Panel A: Determinants and Consequences of Reporting Credibility in the Short-Term



Panel B: Determinants and Consequences of Reporting Credibility in the Long-Term



*** Indicates $p < 0.01$ two-tailed.
^{NS} Indicates $p > 0.10$ two-tailed.
Figure 3 reports the standardized regression coefficients for two path analyses examining (1) the determinants of management's reporting credibility in the short- and long-term, and (2) whether changes in management's reporting credibility affect investors' willingness to rely on management's subsequent disclosures. Participants' cognitive reactions were measured by asking them to indicate the degree to which they attributed management's disclosure decision to management's honesty. Participants' affective reactions were measured by asking participants to assess the direction and strength of their feelings related to the news valence. The change in management's reporting credibility was calculated by subtracting participants' pre-test composite credibility assessments from their post-test composite credibility assessments (see Appendix). Willingness to rely was measured by providing participants with an additional management disclosure following exposure to the experimental manipulations and asking them to indicate their willingness to rely on that disclosure.

participants viewed an additional management disclosure after experiencing the experimental manipulations and indicated their willingness to rely on the disclosure. Panels A and B in Figure 3 show that the link between participants' credibility difference scores and their willingness to rely on management's subsequent disclosure is statistically significant for both Short-term ($\beta = 0.24$, two-tailed $p < 0.01$) and Long-term ($\beta = 0.28$, two-tailed $p < 0.01$) participants. These results suggest that gains and losses in management's reporting credibility indeed influence their ability to communicate persuasively with investors.

V. CONCLUSIONS

In summary, the results suggest that the factors that affect management's reporting credibility differ depending on whether credibility is assessed in the short-term or the long-term. In the short-term, management's level of forthcomingness affects its reporting credibility. That is, forthcoming disclosure temporarily increases reporting credibility and a lack of forthcomingness temporarily decreases reporting credibility. Moreover, negative news disclosure decisions appear to be especially important in determining management's reporting credibility in the short-term. However, the results also suggest that these credibility effects are fleeting. In the long-term, management's reporting credibility is primarily a function of the valence of news disclosed. That is, managers of firms who report positive news are rated as having higher reporting credibility than managers who report negative news, regardless of the managers' forthcomingness about the news. The paper draws on attribution and affect-based reasoning theories from psychology to offer a theoretical framework that predicts and explains each of these results.

The results extend the existing accounting literature in several ways. First, they further our understanding of the determinants of reporting credibility. Despite the fact that management's reporting credibility has a significant impact on management's ability to communicate with investors, we know surprisingly little about how managers develop credibility. Previous studies (Libby and Tan 1999; Tan et al. 2002) show that managers who provide forthcoming disclosures are more credible than managers who are not forthcoming. However, reporting credibility is not the primary focus in these studies, and consequently the authors do not offer a theory for how and when forthcomingness affects reporting credibility. The present paper extends this early work by developing the first theory-based model of how managers' disclosure decisions influence their reporting credibility. Importantly, the model suggests that the credibility effects demonstrated in prior studies only hold in the short-term. Tests of the model confirm this prediction—there appear to be few, if any, long-term credibility benefits to more forthcoming disclosure.

Second, the results may have implications for researchers who use AIMR reports to measure disclosure quality. AIMR reports are based on financial analysts' annual ratings of firms' disclosure policies. These ratings are typically made a significant time after the firms' disclosure decisions. Consequently, the model suggests that AIMR ratings will be based on analysts' affective reactions to the news disclosed instead of their cognitive reactions to the disclosure decisions. That is, the ratings likely reflect firms' recent financial performance rather than disclosure quality. Lang and Lundholm (1993) provide support for this idea. They show that there is a positive relation between a firm's AIMR ratings and its earnings and stock price performance. They argue that their results are due to firms' tendencies to provide more complete disclosures when they are doing well. In other words, Lang and Lundholm (1993) suggest that firm performance affects a firm's actual disclosures, which, in turn, affect AIMR ratings. My model suggests a simpler explanation. AIMR ratings may be a direct result of recent firm performance, independent of any actual disclosure differences across firms. If this is true, then researchers should use caution when using AIMR ratings to measure disclosure quality.

Third, the results may inform the literature on why firms provide earnings warnings. Studies show that firms that provide warnings about negative earnings news are penalized with larger immediate stock price reactions than firms that do not provide such warnings (Kasznik and Lev 1995; Libby and Tan 1999). Despite this apparent penalty for warning investors about negative news, a number of firms continue to provide negative news earnings warnings (Skinner 1994). One proposed explanation for this phenomenon is that the immediate valuation benefits associated with failing to warn are outweighed by the long-term credibility costs of failing to warn (Kasznik and Lev 1995). However, the present study shows that managers who fail to warn about negative news do not take a long-term credibility hit relative to managers who do warn. If managers are aware of the short-term nature of any credibility effects resulting from nonforthcoming disclosures, then long-term credibility concerns cannot explain negative news earnings warnings.

The psychological forces underlying the study's predictions are likely to be common across a broad range of people. Specifically, the study's short-term credibility findings are expected to generalize to all investors who understand managers' disclosure incentives. In other words, I would expect a similar pattern of short-term results for more experienced investors such as financial analysts because these investors have a highly developed understanding of managers' disclosure incentives. However, highly inexperienced investors may not reward managers for forthcoming disclosure if they do not understand that managers sometimes have incentives to withhold information.

The study's long-term credibility findings should generalize to both more- and less-experienced investors, as research in psychology shows that the role of affect in memory is hard-wired (Damasio 1994). In other words, all investors who feel good (bad) when the company reports positive (negative) unexpected earnings news should show a similar pattern of long-term results. Nevertheless, it is possible that investors who are more experienced than my M.B.A. participants develop coping mechanisms that help them avoid the judgment errors implied by my long-term credibility findings. For example, more sophisticated investors may develop better memory structures that allow them to retain their cognitive reactions for longer periods or review disclosures frequently so they reside permanently in a "short-term" world. If true, then the influence of affect on more sophisticated investors' credibility judgments may be less than that observed in my study. However, research in other domains suggests that even the most sophisticated investors are unlikely to escape the pull of affect-based reasoning. Brown and Perry (1994) show that *Fortune* magazine's annual list of most admired companies, which is based on a survey of 8,000 executives, directors, and market analysts, is unduly influenced by a firm's recent financial performance. Consistent with my theoretical framework and results, they find that a firm's recent financial performance creates a "halo" that colors sophisticated investors' judgments of the firm's other attributes (e.g., social responsibility, human resource function, and management quality).

The present study manipulates management's forthcomingness by varying the timeliness of management's disclosures (i.e., whether management provides a warning about unexpected earnings news). However, other factors, such as disclosure completeness and disclosure accuracy, may also affect the perceived forthcomingness of a management disclosure. Identifying these factors, and determining whether and how they interact, is a potentially rich avenue for future research. Likewise, future researchers may wish to consider whether investors respond to *patterns* in a firm's disclosures. For example, do investors regard failures to provide earnings warnings to be less diagnostic of a firm's forthcomingness when the firm rarely provides such warnings?

Finally, future researchers may wish to examine how the credibility effects of disclosure decisions vary when management has a well-developed reporting reputation. My study focuses on the *development* of investors' beliefs about management. That is, I examine how management's disclosure decisions affect its reporting credibility when investors do not have strongly held beliefs about management. The credibility effects identified here may differ when investors have strong beliefs about management. For example, investors who strongly believe that management is credible may interpret a lack of forthcomingness more charitably than investors who harbor suspicions about management's credibility (Ross and Lepper 1980; Koehler 1993). Future investigations of these issues will provide a more complete understanding of the determinants of reporting credibility.

APPENDIX

Management Reporting Credibility Questions

1. I believe that Dentex management is very competent at providing financial disclosures.
2. I believe that Dentex management has little knowledge of the factors involved in providing useful disclosures.
3. I believe that few people are as qualified as Dentex management to provide useful financial disclosures about Dentex.
4. I believe that Dentex management is very trustworthy.
5. I believe that Dentex management is very honest.
6. I believe that Dentex management may not be truthful in their financial disclosures.

Participants evaluated each statement on a seven-point scale with endpoints labeled 1 = Strongly Disagree and 7 = Strongly Agree. Participants evaluated each statement twice—once before viewing the experimental manipulations (Pre-test) and once after viewing the experimental manipulations (Post-test). To minimize response bias, high credibility was sometimes indicated by agreement (i.e., high ratings) and sometimes indicated by disagreement (i.e., low ratings); before analysis, responses were recoded so that higher responses always indicated greater perceived credibility. After this recoding, responses were summed to form a composite measure of management's reporting credibility. Changes in management's reporting credibility were computed by subtracting participants' pre-test composite credibility assessments from their post-test composite credibility assessments.

REFERENCES

Arbuckle, J. 1999. *Amos 4.0* Chicago, IL: Smallwaters.

Axelrod, R. 1984. *The Evolution of Cooperation.* New York, NY: Basic Books.

Barrett, A. 2003. In the credibility penalty box. *BusinessWeek* (April 28): 80.

Brown, B., and S. Perry. 1994. Removing the financial performance halo from *Fortune*'s "most admired" companies. *Academy of Management Journal* 37 (5): 1347–1359.

Brown, L. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (2): 221–241.

Clore, G., N. Schwarz, and M. Conway. 1994. Affective causes and consequences of social information processing. In *Handbook of Social Cognition*, edited by R. Wyer and T. Srull, Vol. 1, 323–417. Hillsdale, IL: Lawrence Erlbaum Associates.

Cullen, T., D. Eisenberg, and S. Steptoe. 2003. Dialing up a departure. *Time* (January 27): 47.

Damasio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain.* New York, NY: Avon.

Dimma, W. 1996. Executive compensation gets its share. *Business Quarterly* 61 (1): 77–80.

Elgin, B. 2003. A nasty surprise from HP. *BusinessWeek* (September 1): 80.

Elliott, W. B., F. Hodge, J. J. Kennedy, and M. Pronk. 2004. When are graduate business students a reasonable proxy for investors? Working paper, University of Illinois Urbana–Champaign, University of Washington, and Tilburg University.

Finucane, M., A. Alkakami, P. Slovic, and S. Johnson. 2000. The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making* 13: 1–17.

Fiske, S., and S. Taylor. 1991. *Social Cognition*. New York, NY: McGraw-Hill.

Giffin, K. 1967. The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin* 68 (2): 104–120.

Healy, P., and K. Palepu. 1995. The challenges of investor communication: The case of CUC International, Inc. *Journal of Financial Economics* 38 (2): 111–140.

Hirst, D. E., L. Koonce, and J. Miller. 1999. The joint effect of management's prior forecast accuracy and the form of its financial forecasts on investor judgments. *Journal of Accounting Research* 37 (Supplement): 1–24.

Hovland, C., I. Janis, and H. Kelley. 1953. *Communication and Persuasion*. New Haven, CT: Yale University Press.

Hughes, J. 1997. How does your pay stack up? *Financial Executive* 13 (3): 24–26.

Jones, E., and K. Davis. 1965. From acts to disposition: The attribution process in person perception. In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, Vol. 2, 219–266. New York, NY: Academic Press.

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70: 113–153.

Kida, T., and J. Smith. 1995. The encoding and retrieval of numerical data for decision making in accounting contexts: Model development. *Accounting, Organizations and Society* 20 (7/8): 585–610.

———, ———, and M. Maletta. 1998. The effects of encoded memory traces for numerical data on accounting decision making. *Accounting, Organizations and Society* 23: 451–466.

Kline, R. 1998. *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.

Koehler, J. 1993. The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes* 56: 28–55.

Lang, M., and R. Lundholm. 1993. Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research* 31 (2): 246–271.

Leathers, D. 1992. *Successful Nonverbal Communications: Principles and Applications*. New York, NY: MacMillan.

Libby, R., and H. Tan. 1999. Analysts' reactions to warnings of negative earnings surprises. *Journal of Accounting Research* 37(2): 415–435.

———, R. Bloomfield, and M. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27: 775–810.

Markman, A., and D. Gentner. 2001. Thinking. *Annual Review of Psychology* 52: 223–247.

McCroskey, J. 1966. Scales for the measurement of ethos. *Speech Monographs* 33: 65–72.

Mellers, B., A. Schwartz, K. Ho, and I. Ritov. 1997. Decision affect theory: How we feel about risky options? *Psychological Science* 8: 423–429.

Mercer, M. 2004. How do investors assess the credibility of management disclosures? *Accounting Horizons* 18 (September): 185–196.

Nunnally, J. 1978. *Psychometric Theory*. 2nd edition. New York, NY: McGraw-Hill.

Olver, L. 2003. Intrawest earnings warning, guidance met with skepticism. *Dow Jones News Service* (May 13).

Rama, M. 2003. Investors punish Electro Scientific after dismal results. *Dow Jones News Service* (March 21).

Robertson, T., and J. Rossiter. 1974. Children and commercial persuasion: An attribution theory analysis. *Journal of Consumer Research* 1: 13–20.

Ross, L., and M. Lepper. 1980. The perseverance of beliefs: Empirical and normative considerations. *New Directions for Methodology of Social and Behavioral Science* 4: 17–36.

Silverman, R. 2002. GE to change its practices of disclosure. *Wall Street Journal* (February 20): A3.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1): 38–60.

Slovic, P., M. Finucane, E. Peters, and D. MacGregor. 2002. The affect heuristic. In *Intuitive Judgment: Heuristics and Biases*, edited by T. Gilovich, D. Griffin, and D. Kahneman. Cambridge, U.K.: Cambridge University Press.

Spiro, R. 1980. Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior* 19 (1): 84–95.

Steinberg, B. 2003. Interpublic may become more open: Effort to offer more data on performance of units is likely to please investors. *Wall Street Journal* (September 10): B4.

Tan, H., R. Libby, and J. Hunton. 2002. Analysts' reactions to earnings preannouncement strategies. *Journal of Accounting Research* 40 (1): 223–246.

Weiner, B. 1985. "Spontaneous" causal thinking. *Psychological Bulletin* 97 (1): 74–84.

Williams, P. 1996. The relation between a prior earnings forecast by management and analyst response to a current management forecast. *The Accounting Review* 71 (1): 103–113.

Ybarra, O. 2002. Naïve causal understanding of valenced behaviors and its implications of social information processing. *Psychological Bulletin* 128 (3): 421–441.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended August 31, 2004. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*:

	Number of Manuscripts	Percent	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	65	16.46	65	16.46
31 ≤ Days ≤ 60	145	36.71	210	53.16
61 ≤ Days ≤ 90	118	29.87	328	83.04
91 ≤ Days ≤ 120	56	14.18	384	97.21
121 ≤ Days	11	2.78	395	100.00
Total	395	100.00		

The mean review time was 60.68 days; the median review time was 59 days.

New Submissions by calendar year (January 1 to December 31)

1992	241
1993	234
1994	231
1995	195
1996	230
1997	195
1998	196
1999	239
2000	260
2001	260*
2002	324
2003	327
2004 Aug 31	210 (8 mths)

* 2001—in addition to the 260 regular submissions, 68 MSs were submitted to the *TAR* Quality of Earnings Conference.

The **acceptance rate** (defined as number of MSs accepted divided by the number of new submissions) over the past 5 years is between 12–15% per year.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of *The Chicago Manual of Style* (14th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be typed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form*. See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says ..."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at taruw@u.washington.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Terry Shevlin, Senior Editor *TAR*, School of Business, Box 353200, University of Washington, Seattle, WA 98195-3200. The submission fee in U.S. funds is $125.00 for members and $200.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review*, *Accounting Horizons*, *Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

APRIL 2005 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The ***Review*** *is published in January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

201 KENT STATE UNIVERSITY, Department of Accounting, invites applications for an Associate Professor with teaching emphasis in Managerial Accounting beginning in Fall 2005. The successful candidate will also serve as a KPMG Faculty Fellow. Candidates must have a high-quality, established research record. A Ph.D. with a major in Accounting or an appropriate related field is required. Professional certification such as CPA or CMA is highly desirable. All candidates must be able to demonstrate evidence of discovery/integration research and of effective undergraduate and graduate teaching. Research will be expected in high-quality, academic refereed journals for reappointment, tenure, and promotion. Participation in teaching doctoral students and serving on and/or supervising their dissertations will be an important part of the faculty member's responsibilities. Review of applications will begin immediately and continue until the position is filled. If you are interested in this position, please contact and/or mail a copy of your *curriculum vitae* to Dr. Pervaiz Alam, Chair, Faculty Search Committee, Department of Accounting, 576 BSA, Kent State University, Kent, OH 44242; Phone: (330) 672-1121; Fax: (330) 672-2548; Email: palam@bsa3.kent.edu.

202 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity, Affirmative Action Employer.

203 INSTITUTO DE EMPRESA, Madrid (Spain), is seeking faculty for tenure-track positions in Accounting beginning Fall 2005. We seek candidates with a Ph.D. or D.B.A. degree in Accounting from an accredited university who have substantial experience in teaching and research. Candidates must be fluent in English. Experience with the case-teaching method and fluency in Spanish is not required but highly valued. Please submit an up-to-date *curriculum vitae,* three reference letters, publication samples, and a statement describing teaching and research interest to Patricia Garcia, Email: Patricia.Garcia@ie.edu.

204 GRAND VALLEY STATE UNIVERSITY, Seidman College of Business, invites applications for Assistant/Associate Professor in Accounting. Tenure-track position is available Fall 2005 and will remain open until filled. Need Ph.D. or D.B.A. (by December 2005) in Accounting or related field. Professional certification is desirable. Cost/Managerial Accounting and Financial Accounting are primary needs; other fields may be considered. Applicant must demonstrate the ability to teach effectively and conduct research. The Department and College are both accredited by the AACSB. Salary is competitive. For consideration, send resume, three references, sample of teaching evaluations, and copies of publications or representative research to Ms. Kathy Goralski, Accounting Search Committee, Department of Accounting and Taxation, Grand Valley State University, 401 West Fulton Street, Grand Rapids, MI 49504, or send email to goralskk@gvsu.edu. Position is subject to availability of funding. Grand Valley State University is an EO/AA Institution.

205 UNIVERSITY OF NORTHERN COLORADO invites applications for a tenure-track Assistant/Associate level position teaching Financial Accounting and at least one other area with a preference for Governmental Accounting. Candidates must have completed all requirements for doctoral degree in Accounting by the position starting date of Fall 2005. This position will be filled as soon as an acceptable candidate is found. Candidates must be committed to excellence in teaching; professional and academic research; and university and community service. The Monfort College of Business undergraduate-only Accounting program is AACSB-accredited. The College is a 2004 recipient of the Malcolm Baldrige National Quality Award. Candidates should submit a cover letter, *curriculum vitae*, evidence of teaching excellence, and three references to Dr. Terri Gutierrez, Accounting Department, Kenneth W. Monfort College of Business, Campus Box 128, Greeley, CO 80639. UNC, located 50 miles north of Denver, is an AA/EO Employer.

206 UNIVERSITY OF CALIFORNIA, LOS ANGELES, The John E. Anderson Graduate School of Management, invites applications for positions in Accounting at both the junior and senior levels beginning July 1, 2005. Applicants must have a Ph.D. or be a Ph.D. candidate and exhibit outstanding research capabilities. *Curriculum vitae* and working papers should be sent to Professor Brett Trueman, The John E. Anderson Graduate School of Management, Box 951481, UCLA, Los Angeles, CA 90095-1481. Under federal law, the University of California may employ only individuals who are legally authorized to work in the United States as established by providing documents specified in the Immigration Reform and Control Act of 1986. UCLA is an Affirmative Action, Equal Opportunity Employer.

207 MANHATTAN COLLEGE, School of Business, an AACSB-accredited school, seeks applicants for a tenure-track position as an Assistant or Associate Professor of Accounting, beginning in August 2005. Candidates must have a Ph.D. or D.B.A. in Accounting as well as evidence of teaching effectiveness and research competence. Responsibilities include teaching undergraduate courses in Accounting, engaging in scholarly research, and participating in campus activities. The Department seeks a candidate willing to teach Taxation, but all qualified candidates are encouraged to apply. Manhattan College is an independent Catholic college founded by the De La Salle Christian Brothers. The College is in the Riverdale section of New York City. Recruiting begins immediately and continues until the position is filled. Salary and rank will be based on qualifications. Please send a letter of application, *curriculum vitae*, and list of three references to Dr. Mary Michel, Chairperson of Accounting, Law and Computer Information Systems, School of Business, Manhattan College, Riverdale, NY 10471 (or email: mary.michel@manhattan.edu). Manhattan College is an Affirmative Action/Equal Opportunity Employer.

208 FERRIS STATE UNIVERSITY invites applications for a tenure-track Accounting position at the rank of Assistant Professor beginning August 2005. Applicants from all areas of Accounting are encouraged to apply. A doctorate in Accounting or Ph.D. in related field from an accredited institution with Accounting certification is required (ABDs will be considered). Candidates should possess an excellent record in teaching, scholarship, and service, as well as a demonstrated capacity to incorporate current teaching technology. Business experience in the field, professional certification, and relevant teaching experience in Financial Accounting and/or Accounting Systems are preferred. Please submit a letter of interest, *curriculum vitae,* names/contact information of three references, and sample copies of student assessment of instruction to JOB CODE AAA-2620, Dr. Mohammed A. El-Saidi, AFES Dept. Head, Ferris State University, 420 Oak St., PRK-150, Big Rapids, MI 49307. For more information about Ferris, please visit our website at http://www.ferris.edu/. Ferris State University is an Equal Opportunity/Affirmative Action Employer.

209 UNIVERSITY OF ARKANSAS invites applications and nominations for the position of Dean of the Sam M. Walton College of Business. The Dean is the chief academic and administrative officer of the College, and reports to the Provost for Academic Affairs. Salary is competitive and commensurate with experience and qualifications. For the position description and information about the University, please visit: Applicants are requested to submit a letter of application, *curriculum vitae,* and five professional references to Dr. G. David Gearhart, Chair of the Search Committee, Vice Chancellor for University Advancement, 416 Administration Building, University of Arkansas, Fayetteville, AR 72701. The selection process will begin March 1, 2005, and continue until the position is filled. Under the state's Freedom of Information Act, after selected candidates are invited to campus, their names are available to the public upon request. The University of Arkansas is an Equal Opportunity, Affirmative Action Institution.

210 UNITED ARAB EMIRATES UNIVERSITY seeks qualified faculty at the levels of Instructor, Assistant Professor, Associate Professor, and Professor for academic year 2005–2006 (which starts in September 2005). Instructors are expected to be holders of a Master's degree; other ranks must hold the terminal degree in the appropriate discipline. The principal responsibility of each faculty is undergraduate instruction. Active research is expected as well. Excellent research facilities and a variety of support are provided. Opportunities are available to teach in the following areas of Accounting: Auditing, Accounting Information Systems, Cost, and Managerial Accounting. Salaries and benefits include: tax-exempt compensation package; housing; furniture and utilities allowances; employer-paid comprehensive health insurance; children's education allowance; annual airline tickets for employee and family; end-of-service gratuity; generous vacation package. If you are interested in joining a professional team dedicated to ensuring growth and development, please visit: http://academics.uaeu.ac.ae/jobs/.

211 UNIVERSITY OF MINNESOTA, Carlson School of Management, invites applications for tenure-track faculty positions at all ranks in Accounting. We offer competitive salary and benefits in a stimulating research environment. All applicants must possess a doctoral degree (or be close to completion), and have a demonstrated commitment to Accounting research and teaching commensurate with rank. Applications will be accepted until positions are filled. Interested candidates should send applications, including *curriculum vitae,* research papers, and at least two references from established scholars to Chair, Search Committee, Department of Accounting, 3-122 Carlson School of Management, University of Minnesota, 321 19th Avenue South, Minneapolis, MN 55455. The University of Minnesota is an Equal Opportunity Educator and Employer.

212 SETON HALL UNIVERSITY, Department of Accounting and Taxation, invites applications for a tenure-track position as Assistant Professor of Accounting beginning September 2005. Applicants must have a Ph.D. in Accounting and should demonstrate evidence of a strong commitment to excellence in teaching and research scholarship. Teaching responsibilities will be primarily in Financial and Managerial Accounting. Salary is competitive. Please send *curriculum vitae* to Professor David S. Gelb, Chair, Department of Accounting and Taxation, Stillman School of Business, Seton Hall University, 400 South Orange Avenue, South Orange, New Jersey 07079-2692. Applications will be accepted until a successful candidate is found.

213 BUENA VISTA UNIVERSITY, Harold Walter Siebens School of Business, invites applications for tenure-track position at the Assistant Professor level to begin Fall 2005. The successful candidate should have an earned doctorate in Accounting, although ABD or appropriate Master's with certification will be considered. Candidates must be dedicated to excellence in teaching, university, and community service and scholarly activity. Teaching in the areas of Financial, Managerial, Auditing, Tax, and Certificate Preparation is expected. BVU is nestled beside a beautiful 3,000-acre lake in the richly diverse northwest Iowa community of Storm Lake. Interested applicants should submit a cover letter, *curriculum vitae,* names and telephone numbers of four references, unofficial transcripts, and evidence of teaching excellence to Dr. Mary Gill, Associate Dean of Faculty, Buena Vista University, 610 West 4th Street, Storm Lake, IA 50588. See complete description at: http://www.bvu.edu/departments/businessservices/hr/jobopenings.asp. Review of applications will begin immediately and continue until the position is filled. Women and minorities are strongly urged to apply. ADA/EOE.

Guo Tai An Information Technology Co., Ltd.

China Stock Market, Bond Market, Banking and Economic Databases for Research

China GTA Company, a data Vendor (CSMAR database) of Wharton Research Data Service (WRDS) (please visit WRDS website for the details http://wrds.wharton.upenn.edu/), has been specialized at developing China stock market, bond market, banking and economic databases for international academic and industry researchers. All GTA's databases are designed to be compatible with international database standards and have been verified by more than 100 professors and experts.

Since 2000, GTA has been serving over 2,000 international scholars and researchers from more than 120 highly reputable educational and professional institutions, such as Prumerica, Merrill Lynch, Barra, China Securities Regulatory Commission, Shanghai Stock Exchange, Yale University, New York University, Wharton School, University of Wisconsin – Madison, University of California at Irvin, University of Reading(UK), Universuty of Pittsburgh, the National University of Singapore, University of Waikato(NZ), Sydney University, University of Hong Kong, Chinese University of Hong Kong, Hong Kong University of Science and Technology, Tsinghua University, Peking University, Fudan University, Shanghai Jiaotong University, Xiamen University, Shanghai University of Finance and Economics, Xi'an Jiaotong University, and Nanjing University. For details, please visit www.chinagtait.com.

CSMAR major research databases (GTA provides also contracted/customized research service)

China Macro-economy

- China Macro-economy Database
- China Industry Research Database

China Stock Market

- CSMAR China Stock Market Trading Database
- China Stock Trade and Quote Research Database
- CSMAR China Stock Market Financial Database—Annual/ Interim/ Quarterly Report
- China Listed Companies' Financial Ratio Research Database
- China Initial Public Offering Research Database
- China Listed Corporate Governance Research Database
- China Listed Companies' Merger & Acquisition, Asset Restructuring Research Database
- China Stock Market Information Disclosure System
- China Stock Market Related Party Transaction Research Database

China Fund Market

- China Securities Investment Fund Research Database
- China Close-end Fund Trade and Quote Research Database

China Bond Market

- China Listed Bond Research Database
- China Listed Convertible Bond Research Database
- China Bond Trade and Quote Research Database

China Futures Market

- China Futures Research Database

Hong Kong Securities Market

- Hong Kong GEM Stock Market Trading Database
- Hong Kong GEM IPO Research Database

Selected Products for securities analysis purposes:

- Daily updating system of China Stock Market & Accounting Research Database (CSMAR): Including financial and trading data updated on a daily basis.
- China ValueLine (CVL) system
- China Securities Information and Analysis System (CSIAS)
- China Securities Excel System (CSXcel)

For free sample databases, manuals, brochures, and more information, please contact us or visit our website:

GTA Information Technology Co., Ltd. http://www.chinagtait.com
Hong Kong, China: 852-27812531 (Tel) 852-27813773(Fax) Shenzhen China: 86-755-83940235 (Tel) 86-755-83940070 (Fax)
Email:marketing@chinagtait.com maysa@chinagtait.com

THE ACCOUNTING REVIEW
ISSN 0001-4826

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFIC

CENTRAL
LIBRARY